DOCUMENT RESUME

ED 434 120 TM 030 081

AUTHOR King, Jason E.; Melancon, Janet G.; Thompson, Bruce
TITLE Score Validation and Theory Elaboration of a Jungian Personality Measure.
PUB DATE 1999-08-20
NOTE 71p.; Paper presented at the Annual Meeting of the American Psychological Association (107th, Boston, MA, August 20-24, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *College Students; Counseling; Factor Structure; Higher Education; Personality Assessment; *Personality Traits; *Psychometrics; *Scores; *Test Construction; Test Items; Validity
IDENTIFIERS Jung (Carl G); Myers Briggs Type Indicator; *Personal Preferences Self Description Quest

ABSTRACT

Jungian measures have proven extremely popular, selling more than three million copies per year for use in various types of counseling. This study investigated the construct validity of scores from an alternative measure of Jungian personality, the Personal Preferences Self-Description Questionnaire (PPSDQ) (B. Thompson, 1996). Forms of the PPSDQ and the Myers-Briggs Type Indicator (I. Myers and M. McCaulley, 1985) were completed by 394 college students. A variety of first- and second-order factor structure models, as well as concurrent validity models, were evaluated using structural equation modeling (SEM). In addition, factor invariance across gender was also evaluated using SEM. The PPSDQ was found to measure four underlying constructs adequately, as hypothesized. Evidence was also found for a higher-order factor underlying these four dimensions. No gender modeling effects were found in this data set. In general, the PPSDQ yielded results comparable to those from the Myers-Briggs measure. (Contains 4 tables, 16 figures, and 45 references.) (SLD)

ppsdq14a.wp1 7/22/99

Score Validation and Theory Elaboration of a Jungian Personality Measure

Jason E. King
Texas A&M University

Janet G. Melancon
Loyola University (LA)

Bruce Thompson
Texas A&M University
and
Baylor College of Medicine

Paper presented at the annual meeting of the American Psychological Association, Boston, August 20, 1999. The senior author may be contacted via e-mail address: "jasonking@tamu.edu". The second author may be contacted via e-mail address: "melancon@loyno.edu". The junior author and related reprints may be accessed through Internet URL: "http://acs.tamu.edu/~bbt6147/".

Score Validation and Theory Elaboration of a Jungian Personality Measure

Abstract

Jungian measures have proven extremely popular, selling more than 3 million copies per year for use in career and marital counseling, as well as in workplace team building and learning styles assessments. The present study investigated the construct validity of scores from an alternative measure of Jungian personality, the Personal Preferences Self-Description Questionnaire (PPSDQ). Forms of the PPSDQ and the Myers-Briggs measure were completed by 394 college students. A variety of first- and second-order factor structure models, as well as concurrent validity models, were evaluated using structural equation modeling (SEM). Additionally, factor invariance across gender was also evaluated using SEM.

Score Validation and Theory Elaboration of a Jungian Personality Measure

The growing recognition that tests are not valid or reliable (instead, scores have these properties to varying degrees) (Thompson, 1992, 1994) has led to the development of methods to establish both validity generalization (Hunter & Schmidt, 1990; Schmidt & Hunter, 1977) and reliability generalization (Vacha-Haase, 1998). The recognition that validity and reliability of scores vary across test administrations leads naturally to explorations of (a) the variability of psychometric coefficients and (b) the factors that do and do not explain or predict that variability.

This view means that establishing validity is a dynamic process in which we apply theory to data to explore validity, but we also simultaneously consider the fit of models to data as evidence bearing upon whether and in what ways theory should be revised or elaborated. Thus, as viewed by Hendrick and Hendrick (1986), "theory building and construct measurement are [invariably] joint bootstrap operations" (p. 393). In a similar vein, Gorsuch (1983) has noted regarding factor analysis that, "A prime use of factor analysis has been in the development of both the operational constructs for an area [theory elaboration] and the operational representatives for the theoretical constructs [score validation]" (p. 350).

Objectives

Measures of normal variation in personality grounded in Jungian theory have been extremely useful in assessing learning styles and in career and other counseling applications. For example, the measure developed by mother and daughter Myers and Briggs (Myers & McCaulley, 1985) "is *the most widely used* personality instrument, with between 1.5 and 2 million persons completing it each year" (Jackson, Parker & Dipboye, 1996, p. 99, emphasis added). As Yabroff (1990) noted, the measure "brought Jung's typology to a high level of practical application" (p. 6). In short, measures of psychological types are among the measures of personality most frequently used in educational and counseling applications (Thompson & Ackerman, 1994).

However, notwithstanding its popularity, the Myers and Briggs' measure has provoked considerable psychometric controversy. Paired articles debating related measurement issues have appeared, for example, in an issue of the <u>Journal of Counseling and Development</u> (Carlson, 1989; Healy, 1989) and also in an issue of <u>Measurement and Evaluation in Counseling and Development</u> (McCaulley, 1991; Merenda, 1991).

The measure has been criticized for the use of a forced-choice or "ipsative" response format, which causes spurious negative correlations among items (Kerlinger, 1986, p. 463). And the measure has been criticized for yielding dichotomized types rather than continuous scores, and for not acknowledging that some people may have relatively neutral preferences on some dimensions. Therefore, an alternative measure of type has been developed by Thompson

(1996b)--the Personal Preferences Self-Description Questionnaire (PPSDQ) (cf. Arnau, Thompson & Rosen, in press; Kier, Melancon & Thompson, 1998; Mittag, in press). As with the Myers and Briggs measure, the PPSDQ yields scores on four dimensions: Extroversion-Introversion (EI), Sensation-iNtuition (SN), Thinking-Feeling (TF), and Judging-Perceiving (JP).

The objectives of the present study were both to explore the validity of PPSDQ scores and to further elaborate a model of personality invoking the Jungian point of view. Specifically, we addressed three research questions:

1. Do PPSDQ scores delineate the expected four-factor (Extroversion-Introversion [EI], Sensation-iNtuition [SN], Thinking-Feeling [TF], and Judging-Perceiving [JP]) Jungian structure?
2. Are PPSDQ scores free of gender bias, as reflected by parameter invariance across gender?
3. When both PPSDQ and Myers-Briggs scores are together jointly analyzed as measuring normal personality, does a single factor emerge in a second-order hierarchical analysis?

These questions were addressed with structural equation modeling techniques (cf. Thompson, in press) with covariance matrices used as the bases for the analyses, for the reasons specified by Cudeck (1989).

Data Source

Instrumentation

Both a form of the PPSDQ (Thompson, 1996b) and the Myers-Briggs' (cf. Myers & McCaulley, 1985) measures were administered. The Myers-Briggs form we used includes 95 scored items that are forced-choice. The PPSDQ version we employed includes 59 items, which are measured on a seven-point Likert scale. Roughly half the PPSDQ items measuring each of the four constructs are reversed in their wording so as to minimize response set influences.

Participants

We collected PPSDQ and Myers-Briggs data from 420 college students enrolled in a private university located in the southern United States. There were more females (n_F=273; 65.0%) than males (n_M=147; 35.0%) in our sample. The mean age of the sample was 23.82 (SD=9.58). Ethnic groups within the sample included: Whites (n=266; 63.3%), African-Americans (n=75; 17.9%), and Hispanics (n=48; 11.4%). This sample was reasonably similar to our various previous samples (cf. Arnau, Thompson & Rosen, in press; Kier, Melancon & Thompson, 1998; Mittag, in press; Thompson & Melancon, 1995), so results should be reasonably comparable across our studies.

We ultimately deleted 22 cases with missing data, and 4 additional cases detected as outliers as regards data normality. There were more females (n_F=253; 64.2%) than males (n_M=141; 35.8%) in our final sample of 394 participants. The mean age of the final sample was 24.01 (SD=9.10).

Analytic Requirements

Univariate Normality

Several requirements must be met before maximum likelihood theory should be used as a parameter estimation method. First, the data should be distributed multivariate normal (Ashcraft, 1998;

Henson, in press). In assessing this property, a necessary but not sufficient condition is univariate normality. For all items considered individually, skewness coefficients ranged from -1.013 to 1.301, and kurtosis coefficients from -1.079 to 1.506. Thus, the data were slightly non-normal.

We attempted to remedy the problem through the use of item parcels. Item "parcels" or "testlets" are created by combining items in one way or another. It has long been recognized that item data can be combined so as to optimize the normality of data (e.g., Cattell, 1956; Cattell & Burdsal, 1975; Gorsuch, 1983, pp. 294-295). For example, item "testlets" can be created by pairing item responses with opposite skewness (e.g., create parcel 1 on a scale by adding the scores on the most negatively skewed item to the scores on the most positively skewed item within a given scale).

Combining items into "parcels" also results in more parsimonious model tests. One feature of this parsimony is that the rank of the estimated matrix of associations can be radically reduced. For example, if 78 items were the basis of analyses, initially 78 variances and 3003 (78 x 77 / 2 = 6006 / 2) unique covariances are estimated, and then the parameters to reproduce these coefficients are estimated. If the same 78 item responses are aggregated only into scores on 36 "doublets," initially only 36 variances and 630 (36 x 35 / 2 = 1260 / 2) unique covariances are estimated, and then the parameters to reproduce these coefficients are estimated.

The number of model parameters is also itself reduced by this process. Fitting more parsimonious models to reproduce fewer estimated population values in the matrix of associations leaves less room for sampling error to impact the estimation process. This in turn theoretically leads to results that better generalize.

It was decided to aggregate the individual items into parcels for two reasons. First, while the data did not depart substantially from univariate normality, mild departures can compound in the multivariate factor space and result in appreciable multivariate non-normality. Second, it has been suggested that one have five cases for every freed parameter in a given model (Bollen, 1989). In testing a model involving 59 items, this requirement is not met with a sample size of 394 unless only one parameter per item is estimated. In this case, the potential advantage to using item parcels is the ability to obtain more valid model tests and estimates, given the small-to-moderate sample size and the relative non-normality of this data set. The primary disadvantage is the loss of interpretability at the item level.

Two sets of item parcels were constructed. Under the first method, items were paired based on the magnitudes and signs of their skewness coefficients. Items skewed negatively were matched with those skewed positively in an effort to offset the effect of skewness on the data. Three to five parcels were created per hypothesized dimension (e.g., four EI sublets were created). The 16 parcels yielded a mean skewness coefficient of -.048 (SD = .234), ranging from -.331 to .556. These values are superior to those obtained using individual items and more likely to be distributed as multivariate normal.

The second method we used to obtain parcels entailed exploratory factor analysis of the items hypothesized to saturate each construct. An identical number of parcels (16) were constructed using this method. These parcels yielded an average skewness coefficient of -.108 (SD = .356), ranging from -.813 to .520.

Multivariate Normality

To assess the sufficient condition (multivariate normality), after ordering cases by the Mahalanobis distance of each case's set of scores from the variable centroids, the distance for each participant was plotted with the expected chi-square value associated with the individual's position in the distribution of distance scores. Computer program MULTINOR (Thompson, 1990) was used to acquire the graph. This procedure was employed in addition to statistical significance tests, due to the inherent limitations of statistical tests (Schmidt & Hunter, 1997; Thompson, 1996a, 1999).

Using the present graphical method, one can identify not only the individuals contributing to non-normality, but also obtain a relative index of degree of normality. Four outliers, identified as such by the multivariate plot, were removed. These four participants were classified as extreme when using either the factor analytic or skewness parcels. For the item parcels, perfect multivariate normality could not be assumed since all coordinates did not fall along a straight line, as reported in Figure 1; however, the departure did not appear to be extreme.

INSERT FIGURE 1 ABOUT HERE.

Mardia's coefficient of multivariate kurtosis was also computed for the data. For the individual items (i.e., before parceling), Mardia's coefficient equalled 535.511 (critical ratio = 61.929). These values indicate a large degree of non-normality in the distribution. For the *factor analytic* item parcels, Mardia's coefficient equalled 61.721 (c.r. = 25.524); for the *skewness* parcels, Mardia's coefficient was 44.208 (c.r. = 18.282). Though the data were still non-normal, the degree of non-normality diminished substantially from the original items to the factor analytic parcels to the skewness parcels. Thus, the *skewness parcels* were used for the primary analyses in the present study.

Because the data were to be partitioned by sex to evaluate the invariance of estimates across gender, each group's distribution was tested for multivariate normality as well, as reported in Figure 2. The males' parcel scores were more non-normal (Mardia's coefficient = 49.704, c.r. = 12.298) than were the females' scores (Mardia's coefficient = 31.990, c.r. = 10.601).

INSERT FIGURE 2 ABOUT HERE.

Sample Size

A second requirement for ML estimation is a large sample size (Thompson, in press). The parameter estimates and fit indices are only assumed to be valid asymptotically. As explained earlier, item parcels were created to meet the suggested five-cases-per-freed-parameter guideline (Bollen, 1989). This requirement was met for all models involving item packets, given the present sample of 394 valid cases. The largest number of freed parameters to be estimated was 76 for the model testing the invariance of factor structure across gender. To meet Bollen's criteria for this model, 380 data points would be required to estimate 76 parameters. However, for models involving individual items, well over 100 parameters were estimated (126 for the correlated factors model). Consequently, results from analyses at the item level should probably not be interpreted. Similarly, results from the equal covariances model (137 parameters estimated) should be viewed skeptically as well.

Methodological Issues

Computer Software

AMOS 3.6 was used for all analyses. Results obtained from AMOS should mirror those obtained using LISREL or EQS. According to Cox (1995), almost all estimates will be identical through two decimal places.

Scaling of Latent Factors

For the initial analyses, the scales of latent factors were set by constraining the factor variances to unity. When testing higher-order factors and the invariance of parameters across groups, the scale was set using indicators. To choose which measured variables would have their paths to constructs set to unity for model identification purposes, alpha-if-deleted statistics were computed for the complete set of item parcels for a given scale. For each of the 4 Jungian (e.g., EI) constructs, the item parcel for which alpha deteriorated the most if the parcel was not used in computing score reliability for the complete parcel set was selected to scale the latent factors, because scores on this parcel appeared to contribute the most to construct reliability (Byrne, 1989, 1994).

Estimation Method

Maximum likelihood (ML) estimation maximizes the fit between the estimated population variance/covariance matrix and that implied by the model. While generalized least squares (GLS) solutions were also obtained for most models, we interpreted the ML estimates. Although some suggest using GLS for estimating structural models, in simulation studies Chou and Bentler (1995) found that ML estimates reject true parameter values more consistently than either GLS or ADF (asymptotic distribution free) methods (see their Table 3.5 on p. 53). The authors stated, "All the fit indices obtained from ML performed much better than those obtained from GLS and ADF and should be preferred indicators" (p. 94). It seems that ML is superior to the other two theories, at least when the data are reasonably multivariate normal.

Measures of Model Fit

Following the recommendations of various methodologists (cf. Fan, Thompson & Wang, 1999; Hoyle & Panter, 1995; Hu & Bentler, 1995, 1999) the chi-square statistic, df, and p-value; the GFI absolute fit index; and the TLI (NNFI), IFI, and CFI relative fit indices were all reported. Additionally, the chi-square/df ratio along with the RMR, RMSEA, and AIC absolute fit indices were also included. However, there are problems with such indices. McDonald (1997) cautioned:

> I do not believe that we presently know how to use such indices and in particular I do not believe on current evidence that global indices of approximation can or should be used as the sole basis for a decision that a restrictive model is acceptable. (p. 217)

The chi-square statistics were chosen due to the extreme popularity of such measures in SEM, along with their utility and validity, when interpreted properly. While statistical significance tests usually assume a "nil null" (Cohen, 1994), nested models do not. The change in chi-square actually reflects the difference between two plausible models. However, since the probability of the test statistic is still largely affected by sample size (cf. Thompson, 1996a, 1999), other indices are needed as well.

The GFI indexes the relative amount of the observed variance/covariance matrix accounted for by the implied model, and are analogous to an $\underline{R}^2$ statistic. Hu and Bentler (1995) noted that "Marsh et al. (1988) found that GFI appeared to perform better than any other absolute index (e.g., AGFI, CAK, CN, RMR, etc.)" (p. 91).

The root mean square residual (RMR) is the average of the fitted residuals obtained from subtracting the implied model variance/covariance matrix from the observed. Hu and Bentler (1995) suggested always including the standardized RMR. Becausee Amos does not report the standardized RMR, interpretation is more difficult. In general, lower values are to be preferred, with 0 indicating a perfect fit to the sample data.

The root mean square error of approximation (RMSEA) measures the lack of fit per degree of freedom. It indicates the potential fit of the model to the population parameters. Values below .05 are considered to be a "close fit", and below .08 "reasonable" (Browne & Cudeck, 1993, p. 144). This measure was included due to the "strong urgings" of MacCallum to include such an index which penalizes for model complexity (1995, p. 30).

The Akaike Information Criterion (AIC) was primarily reported to compare non-nested models. This absolute index also penalizes for increasing the number of parameters being estimated.

Type-2 relative fit indices are useful for comparing models but do not measure explained variance. The Tucker-Lewis Index (TLI) denotes the relative improvement in fit per degree of freedom for a given model compared with a baseline model. The incremental fit index (IFI) is similar but is more consistent across estimators. The only Type-3 relative index included here was the comparative fit index (CFI). It first replaces the central chi-square with a noncentral chi-square and then measures the relative reduction in lack of fit. These Type-2 and -3 indices make use of more information, but the assumed distributions (e.g., the noncentral chi-square) may be wrong (Hu & Bentler, 1995).

In general, accepting models based only on obtaining "the 'magic' .90 level" (Judd, Jessor, & Donovan, 1986) was avoided. Hu and Bentler (1995) stated that such a standard is "clearly an inadequate rule" and that "we are hardly able to point to a condition for which it yields appropriate results" (p. 95).

Measurement Models Using All Items

Uncorrelated Factors

As mentioned earlier, due to the relatively small sample size present, one should not interpret parameter estimates or fit indices at the item level. The models reported in this section are included only for the sake of completeness and possible use in future research. As shown in Figure 3, the 59 items were hypothesized to reflect four underlying factors. While all solutions were proper, the uncorrelated factors model fit the data extremely poorly, $\chi^2 = 4297.98$, $p < .001$, GFI = .71, as reported in Table 1.

INSERT FIGURE 3 AND TABLE 1 ABOUT HERE.

The primary reason for the large lack of fit probably stems from the number of degrees of freedom in the model, 1710, relative to the number of parameters estimated, 120. The data do suggest an underlying factor structure, as reflected in the comparison with an independence model specifying no covariances among variables, change $\chi^2 = 4846.93$, $p < .001$. This relatively uninteresting finding simply validates our treatment of the variables as being related in some way.

Correlated Factors

Allowing the four latent variables to correlate increased the fit marginally, change $\chi^2 = 354.24$, $p < .001$, change GFI = .02, as reported in Table 1. None of the comparative fit indices are large, nor is the average RMR value small, as would be hoped. But given the large _df_ in this model as well, one would need to take the small number of parameters being estimated into account. The χ^2 / _df_ ratio was 2.31. This ratio suggests a fair fit to the data (though one would have to reject this model based on all other indices).

All of the critical ratios for regression weights were larger than 2.0, and most were above 4.0, suggesting that the items are important in measuring each construct. The EI construct was negatively correlated with the other three latent variables, while SI and JP yielded an r of .801. Further, variance-accounted-for indices indicated that some items were not being measured well by the underlying factors (r^2 for EI31 as low as .032). Future assessments of the measurement model underlying the individual PPSDQ items should be carried out with a much larger sample size. More complex analyses at the item level are not discussed here due to the sample size limitation.

Measurement Models Using Item Parcels

Factor Analytic Parcels

Almost all fit indices and fit statistics reflected a slightly poorer fit in the factor analytic parcels when compared with the skewness parcels, as reported in Tables 2 and 3. This was probably due to the more severe non-normality of the factor analytic parcels, described previously. As noted above, Chou and Bentler (1995) found that when the data are multivariate normal ML outperformed GLS and ADF methods not only in estimating parameters but also in providing more accurate fit statistics. Because the factor analytic parcels were less normally distributed, one would expect these fit statistics to be somewhat more inaccurate (although one cannot know whether the reduced fit was due to non-normality or to an actual poorer fit with these testlets). Consequently, here we interpret primarily the results obtained using the skewness parcels.

INSERT TABLES 2 AND 3 ABOUT HERE.

Uncorrelated Factors

The uncorrelated factors model did not result in an acceptable fit for the skewness parcels, $\chi^2 = 625.35$, $p < .001$, χ^2 / $df = 6.01$, GFI = .83, as reported in Table 3 and Figure 4. [Note that the values in the graph beside the observed variables are NOT the error variances, but here are the percentages of variance explained by the model for each observed variable.] All squared multiple correlations were above .40 excepting four parcels, three of which were in the TF factor. These results, coupled with a modification index of 159.665 for JP and SI, indicated a potential cross-loading problem.

INSERT FIGURE 4 ABOUT HERE.

Correlated Factors

The four correlated factors model (see Figure 5) was the first model to fit the sample data well. Though the test statistic was statistically significant, $\chi^2 = 302.66$, $p < .001$, the absolute fit index GFI indicated that 91% of the variability within the variance/covariance matrix was being accounted for by the implied model. All three relative fit indexes were above .91. Further, the RMSEA was below .08, indicating a potentially "acceptable" fit to the data.

INSERT FIGURE 5 ABOUT HERE.

From inspecting the standardized residual covariances, it appeared that one of the SI parcels was correlated with three of the TF parcels. Cross-loading the parcel onto the TF factor increased the fit marginally (GFI increased from .91 to .917; chi-square change ≈ 20), but the added complexity in interpretation and calculation of scale scores would seem to argue against this slight and atheoretical increment in fit. Given that this was one of the two models theorized to underlie the population data (the other including a higher-order factor), these results indicated a relatively good fit between the theory and the data.

Other Potential Models

A more restrictive model ($df = 104$) is the one factor model. Here, one general factor is assumed to underlie the covariances among variables instead of four. The fit of this model, presented in Figure 6, was poor. As reported in Table 3, $\chi^2 = 1362.44$, $p < .001$, GFI = .64, and the model was abandoned as a viable alternative to the four factor hypothesis.

INSERT FIGURE 6 ABOUT HERE.

For the Myers-Briggs data only, a just-identified model was estimated specifying four factors. This model (not portrayed here) yielded a correlation coefficient between SI and JP of .492. The correlated factors model using the PPSDQ data resulted in an even higher correlation between the two factors ($\underline{r}$ = .80). Because of these high correlations, a possible three-factor PPSDQ model was hypothesized such that SI and JP items saturated a common single factor, as reported in Figure 7. This model fit the data reasonably well (GFI = .89, CFI = .90), but not as well as a four factor solution (change χ^2 = 65.02, $\underline{p}$ < .001). Even in the three-factor solution, EI continued to correlate negatively with every other factor in the analysis ($\underline{r}$ = -.30 with SiJp; $\underline{r}$ = -.28 with TF).

INSERT FIGURE 7 ABOUT HERE.

Models Nested Within the Four Factor Solution

Since four correlated factors fit the data better than one, three correlated, or four uncorrelated factors, this model was chosen as best. But it still remained to be seen whether there are higher-order factors that underlie the first-order latent variables. Since the fit of first- and second-order factor analyses will be very similar, specifying a higher-order model versus correlated first-order factors should be based primarily on theory (Byrne, 1994, p. 118). Theory suggests that a global personality construct should underlie the psychological types.

A less restrictive model involving two higher-order factors was first tested, as reported in Figure 8. This model fit the data adequately, GFI = .91, RMSEA = .07.

INSERT FIGURE 8 ABOUT HERE.

In comparing the Figure 8 two higher-order factors solution to the only one higher-order factor model, as reported in Figure 9, results again indicated no substantial differences between the models (GFI = .91 for both, CFI = .92 for both), in spite of the statistically significant test statistic, change χ^2 = 5.58, $\underline{p}$ = .018. However, the percentage of variance explained differed across solutions (cf. Figures 8 and 9). The two factor solution increased the EI $\underline{r}^2$ from .11 to .17, and boosted the TF $\underline{r}^2$ 20%. This probably occurred due to the large correlation reported earlier ($\underline{r}$ = .80) between SI and JP. The addition of another factor freed the EI and TF factors to load elsewhere. [Figure 10 presents a similar analysis in which all 4 paths from the first-order factors to the second-order factor were constrained to equal unity.]

INSERT FIGURES 9 AND 10 ABOUT HERE.

For this reason, one might argue for a two higher-order factors solution. But one could also rationally maintain a one higher-order factor solution based on parsimony and the absence of any substantive changes in fit indices when comparing the two models. The more parsimonious option was selected here for the reasons just stated and given the non-statistically significant results obtained when a GLS solution was considered (change χ^2 = 2.059, p = .151). In our opinion, the relatively slight evidence favoring a two factor model does not outweigh the more parsimonious and theoretically meaningful one higher-order factor model.

The remainder of the analyses in this paper address results for the two models that appear to be the most reasonable (i.e., the four correlated first-order factors model and the one higher-order factor model). At this point one would usually assess the indirect effects of the higher-order factor on each observed variable. But given the aggregated data used here (i.e., item parcels), such findings would seem to convey little information.

Between Group Comparisons

Test of Equal Covariances

The first between group comparison we assessed was the comparability of variance/covariance matrices across gender. Results indicated relative equality between female and male covariance matrices, χ^2 = 161.23, p = .061, GFI = .95, RMSEA = .02, as reported in Table 4. These findings suggest that models constraining parameters for both groups to be invariant may be viable.

INSERT TABLE 4 ABOUT HERE.

Tests Assuming Four Correlated First-Order Factors

Table 4 presents the estimates of the similarity of measurement models for female and males. This model, constraining parcels to saturate identical factors across groups, fit the data reasonably well, controlling for the many degrees of freedom, χ^2 / df = 2.00, GFI = .89, RMSEA = .05, CFI = .92. The large chi-square value (391.92, p < .001) was obtained by adding the chi-square for females (χ^2 = 232.02) with that for males (χ^2 = 159.84), as reported in Table 5.

INSERT TABLE 5 ABOUT HERE.

In this case, the overall chi-square was influenced more by the lack of fit in the females model (59% of total chi-square) than the males (41%). However, the divergence in percentages is not surprising, given that female sample size was almost twice that of males, as sample size does itself inflate statistical significance tests, even holding model fit completely constant.

Second, we tested the equality of factor pattern (lambda) coefficients for both groups on the four factor solution. The fit of this model was comparable to one in which the pattern coefficients could vary across groups (change χ^2 = 18.48, p = .102). Next, the error terms associated with each manifest variable were constrained to be equal across gender. Again the differences were not statistically significant (change χ^2 = 19.39, p = .249). Finally, the correlations among the four latent variables were constrained to be equal across gender, in addition to the previous constraints imposed. Once more this model fit the data roughly as well as one allowing all three sets of parameters to vary across groups (change $\chi^2 = 11.78$, $p = .067$).

These persuasive results indicate that the factor structure (assuming four first-order variables) for this data set is similar for both genders. Final models for each group are presented in Figures 11 and 12.

INSERT FIGURES 11 AND 12 ABOUT HERE.

Tests Assuming One Second-Order Factor

Because we selected the higher-order factor model as best fitting and most parsimonious, we next evaluated the invariance of female and male estimates for this model. We ultimately tested whether the measurement model underlying both groups contained a single higher-order factor.

First, we compared the similarity of measurement models across groups, assuming one higher-order factor. Once degrees of freedom were accounted for, this model fit well, χ^2 / df = 2.02, GFI = .89, RMSEA = .05, CFI = .92. As in previous comparisons, the overall chi-square was influenced more by lack of fit in the females distribution ($\chi^2 = 241.01$, representing 60% of variation in the overall chi-square) than in the males ($\chi^2 = 162.28$, representing 40% of total), as reported in Table 5.

Constraining pattern coefficients to be identical across groups did not result in a poorer fitting model (change $\chi^2 = 16.93$, $p = .152$; most fit indices were identical). Assuming equal error variances was tenable (change $\chi^2 = 19.13$, $p = .262$), as was assuming equal pattern coefficients for the first-order factors on the higher-order factor (change $\chi^2 = 9.46$, $p = .051$). Again, these results indicate similar factor structure across groups.

Testing the Assumption of a Higher-Order Factor

The final models estimated for each group are depicted in Figures 13 and 14. One reason for the slightly poorer fit for the females involved the EI items. Less factor variance was explained for the females (8%) than the males (11%). This can also be seen in the smaller pattern coefficient (-.29 versus -.33). Further, the TF factor was weighted stronger for the females (.56) than the males (.50), suggesting that TF, SI, and JP tended to "clump together" in the female data, while excluding EI. For the males, all four factors were more equally related.

INSERT FIGURES 13 AND 14 ABOUT HERE.

Between Measure Comparisons

Having found two models that fit the data adequately for both groups of participants (i.e., four correlated first-order factors or a one higher-order factor model), we next decided to compare PPSDQ results with scores on the Myers-Briggs. Initially, both variables for each Myers-Briggs factor (e.g., Extroversion score and Introversion score) were entered into an analysis with PPSDQ item parcels. This resulted in the Myers-Briggs items explaining most of the variance in the model because the two Myers-Briggs scale scores for each factor are obtained by summing items purported to measure the same construct but scaled in opposite directions (e.g., all items measuring Extroversion are summed, as are all items measuring Introversion). Their mutual occurrence in a factor analysis distorts results due to the strong linear relationship between each scale score within each factor. Thus, here a Myers-Briggs composite was created by averaging the two scores on each type (e.g., Extroversion scores were reverse coded and then averaged with Introversion scores).

Figure 15 depicts the four correlated factors model for the two combined measures. [Note that each Myers-Briggs composite is the far right indicator under each factor.] This model fit the data fairly well, GFI = .89, RMSEA = .07, CFI = .92, as reported in Table 6. However, the test statistic was large, $\chi^2 = 482.27$, $p < .001$. Also, the χ^2 / df ratio (2.94) was not as small as desired. On inspecting the factor pattern coefficients for the item packets and Myers-Briggs variables, the two were generally comparable. Most r^2's were above .5 indicating an adequate percentage of variance explained for each observed variable (Byrne, 1989). Exceptions were several of the TF item parcels, as was discovered in earlier analyses.

INSERT FIGURE 15 AND TABLE 6 ABOUT HERE.

A higher-order factor model, reported in Figure 16, fit the data as well as a four correlated factors model (change $\chi^2 = 4.62$, $p = .099$). Again the higher-order factor was being dominated by SI ($\lambda = .95$) and JP ($\lambda = .82$). In general, both PPSDQ and Myers-Briggs scores appeared to be measuring the same constructs. Of course, the measures were not perfectly correlated, but this would not be desired if the PPSDQ were hypothesized to be an improvement over the older Jungian measure.

INSERT FIGURE 16 ABOUT HERE.

Conclusions

Several conclusions can be drawn from these results. First, the PPSDQ was found to adequately measure four underlying constructs, as hypothesized. Second, there is evidence suggesting that a higher-order factor might underlie the four dimensions. This construct could be considered a general personality factor. Third, interpretive results can be generalized to both females and males; there were no gender moderating effects present in this data set. Finally, while not measuring the constructs in exactly the same manner, the PPSDQ yields results comparable to those obtained from the Myers-Briggs measure.

Measures of Jungian type are among the most commonly used measures of normal personality variations across diverse applications, including learning styles assessment and guidance counseling. However, the Myers-Briggs measure has been criticized on the various psychometric grounds summarized previously. The present results together with prior results (cf. Arnau, Thompson & Rosen, in press; Kier, Melancon & Thompson, 1998; Mittag, in press) suggest that the PPSDQ may be used for the same important purposes, while at the same time avoiding various pitfalls associated with the alternative measure.

References

Arnau, R.C., Thompson, B., & Rosen, D.H. (in press). Alternative measures of Jungian personality constructs. Measurement and Evaluation in Counseling and Development.

Ashcraft, A.S. (1998, April). Ways to evaluate the assumption of multivariate normality. Paper presented at the annual meeting of the Southwestern Psychological Association, New Orleans. (ERIC Document Reproduction Service No. ED 418 095)

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1989). A primer of LISREL: Basic applications and programming for confirmatory factor analytic models. New York: Springer-Verlag.

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming. Thousand Oaks: Sage.

Carlson, J. G. (1989). Affirmative: In support of researching the Myers-Briggs Type Indicator. Journal of Counseling and Development, 67, 484-486.

Cattell, R. B. (1956). Validation and intensification of the sixteen personality factor questionnaire. Journal of Clinical Psychology, 12, 205-214.

Cattell, R. B., & Burdsal, C. A. (1975). The radical parcel double factoring design: A solution to the item-vs-parcel controversy. Multivariate Behavioral Research, 10, 165-179.

Chou, C.P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 37-55). Thousand Oaks, CA: Sage.

Cohen, J. (1994). The earth is round ($p < .05$). American Psychologist, 49, 997-1003.

Cox, J. J. (1995). Amos, EQS, and LISREL for Windows: A comparative review. Structural Equation Modeling, 2(1), 79-91.

Cudeck, R. (1989). The analysis of correlation matrices using covariance structure models. Psychological Bulletin, 105, 317-327.

Fan, X., Thompson, B., & Wang, L. (1999). The effects of sample size, estimation methods, and model specification on SEM fit indices. Structural Equation Modeling, 6, 56-83.

Gorsuch, R.L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Healy, C. C. (1989). Negative: The MBTI: Not ready for routine use in counseling. Journal of Counseling and Development, 67, 487-488.

Hendrick, C., & Hendrick, S. (1986). A theory and method of love. Journal of Personality and Social Psychology, 50, 392-402.

Henson, R.K. (in press). Multivariate normality: What is it and how is it assessed?. In B. Thompson (Ed.), Advances in social science methodology (Vol. 5). Stamford, CT: JAI Press.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 158-176). Thousand Oaks: Sage.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 76-99). Thousand Oaks: Sage.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1-55.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage.

Jackson, S. L., Parker, C. P., & Dipboye, R. L. (1996). A comparison of competing models underlying responses to the Myers-Briggs Type Indicator. Journal of Career Assessment, 4, 99-115.

Judd, C.M., Jessor, R., & Donovan, J.E. (1986). Structural equation models and personality research. Journal of Personality, 54, 149-198.

Kerlinger, F.N. (1986). Foundations of behavioral research (3rd ed.). New York: Holt, Rinehart and Winston.

Kier, F. J., Melancon, J.G ., & Thompson, B. (1998). Reliability and validity of scores on the *Personal Preferences Self-Description Questionnaire (PPSDQ)*. Educational and Psychological Measurement, 58, 612-622.

MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 158-176). Thousand Oaks: Sage.

McCaulley, M. H. (1991). Additional comments regarding the Myers-Briggs Type Indicator: A response to comments. Measurement and Evaluation in Counseling and Development, 23, 182-185.

McDonald, R. P. (1997). Goodness of approximation in the linear model. In J. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), What if there were no significance tests? (pp. 199-219). Mahwah, NJ: Erlbaum.

Merenda, P. F. (1991). Additional comments regarding the Myers-Briggs Type Indicator. Measurement and Evaluation in Counseling and Development, 23, 179-181.

Mittag, K. (in press). Measuring the Jungian personality types of high school students. Assessment.

Myers, I. B., & McCaulley, M. H. (1985). Manual: A guide to the development and use of the Myers-Briggs Type Indicator. Palo Alto, CA: Consulting Psychologists Press.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. Journal of Applied Psychology, 62, 529-540.

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In J. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), What if there were no significance tests? (pp. 37-64). Mahwah, NJ: Erlbaum.

Thompson, B. (1990). MULTINOR: A FORTRAN program that assists in evaluating multivariate normality. Educational and Psychological Measurement, 50, 845-848.

Thompson, B. (1992). Two and one-half decades of leadership in measurement and evaluation. Journal of Counseling and Development, 70, 434-438.

Thompson, B. (1994). Guidelines for authors. Educational and Psychological Measurement, 54, 837-847.

Thompson, B. (1996a). AERA editorial policies regarding statistical significance testing: Three suggested reforms. Educational Researcher, 25(2), 26-30.

Thompson, B. (1996b). Personal Preferences Self-Description Questionnaire. College Station, TX: Psychometrics Group.

Thompson, B. (1999). If statistical significance tests are broken/misused, what practices should supplement or replace them?. Theory & Psychology, 9, 165-181.

Thompson, B. (in press). Ten commandments of structural equation modeling. In L. Grimm & P. Yarnold (Eds.), Reading and understanding multivariate statistics (Vol. 2). Washington, DC: American Psychological Association. (Originally an Invited paper presented at the 1998 annual meeting of the U.S. Department of Education, Office of Special Education Programs (OSEP) Project Directors' Conference, Washington, DC; ERIC Document Reproduction Service No. ED 420 154)

Thompson, B., & Ackerman, C. (1994). Review of the Myers-Briggs Type Indicator. In J. Kapes, M. Mastie, & E. Whitfield (Eds.), A counselor's guide to career assessment instruments (3rd ed., pp. 283-287). Alexandria, VA: American Counseling Association.

Thompson, B., & Melancon, J. (1995, January). Measurement integrity of scores from a self-description checklist evaluating Myers-Briggs Type Indicator (MBTI) types: A confirmatory factor analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas. (ERIC Document Reproduction Service No. ED 380 487)

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. Educational and Psychological Measurement, 58, 6-20.

Yabroff, W. (1990). The inner image: A resource for type development. Palo Alto, CA: Consulting Psychologists Press.

Table 1
Comparison of Models Containing All Items

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Independence	9144.91	1770	<.001	5.17	--	--	--	.33	.44	.10	9264.91	.00	.00	.00
Uncorr'd Factors	4297.98	1710	<.001	2.51	4846.92	60	<.001	.71	.30	.06	4537.98	.65	.64	.65
Corr'd Factors	3943.74	1704	<.001	2.31	354.24	6	<.001	.73	.18	.06	4195.74	.70	.69	.70

Table 2
Comparison of Models Containing Factor Analysis Item Parcels

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Independence	2205.17	120	<.001	18.38	--	--	--	.46	5.36	.21	2237.17	.00	.00	.00
Uncorr'd Factors	722.12	104	<.001	6.94	1483.05	16	<.001	.82	3.77	.12	786.12	.71	.66	.70
Corr'd Factors	396.08	98	<.001	4.04	326.04	6	<.001	.89	1.32	.09	472.08	.86	.83	.86

Table 3
Comparison of Models Containing Skewness Item Parcels

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Independence	2761.78	120	<.001	23.01	--	--	--	.39	4.33	.24	2793.78	.00	.00	.00
Uncorr'd Factors	625.35	104	<.001	6.01	2136.43	16	<.001	.83	2.87	.11	689.35	.80	.77	.80
Corr'd Factors	302.66	98	<.001	3.09	322.69	6	<.001	.91	.85	.07	378.66	.92	.91	.92
Models compared to Four Corr'd Factors														
One General Factor	1362.44	104	<.001	13.10	1059.78	6	<.001	.64	2.13	.18	1426.44	.53	.45	.52
Three Factors (Corr'd)	367.68	101	<.001	3.64	65.02	3	<.001	.89	.90	.08	437.68	.90	.88	.90
Comparisons of Nested Models														
Four Corr'd Factors	302.66	98	<.001	3.09	--	--	--	.91	.85	.07	378.66	.92	.91	.92
Two Higher-Order Factor	306.98	99	<.001	3.10	4.33	1	.037	.91	.87	.07	380.98	.92	.91	.92
One Higher-Order Factor	312.56	100	<.001	3.13	5.58	1	.018	.91	.90	.07	384.56	.92	.90	.92
One Higher-Order Factor (constraining lambdas = 1.0)	575.18	104	<.001	5.53	262.62	4	<.001	.83	2.94	.11	639.18	.82	.79	.82

Table 4
Assessment of Factorial Invariance Across Sex

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Equal Covariances	161.23	135	.061	1.19	--	--	--	.95	1.03	.02	435.23	.99	.98	.99
Identical Measurement Model (Four Corr'd Factors)	391.92	196	<.001	2.00	--	--	--	.89	.93	.05	543.92	.92	.91	.92
+ Equal Lambdas	410.41	208	<.001	1.97	18.48	12	.102	.89	1.01	.05	538.41	.92	.91	.92
+ Equal Thetas	429.79	224	<.001	1.92	19.39	16	.249	.88	1.00	.05	525.79	.92	.91	.92
+ Equal Phis	441.58	230	<.001	1.92	11.78	6	.067	.88	1.17	.05	525.58	.92	.91	.92
One Higher-Order Factor	403.35	200	<.001	2.02	11.43	4	.022	.89	.95	.05	547.35	.92	.90	.92
+ Equal Lambdas	420.28	212	<.001	1.98	16.93	12	.152	.88	1.03	.05	540.28	.92	.91	.92
+ Equal Thetas	439.41	228	<.001	1.93	19.13	16	.262	.88	1.03	.05	527.41	.92	.91	.92
+ Equal 2nd Order Lambdas	448.87	232	<.001	1.93	9.46	4	.051	.88	1.19	.05	528.87	.92	.91	.92

Table 5
Measurement Models by Sex

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Females														
Independence	1792.17	120	0	14.93	--	--	--	.404	4.10	.235	1824.17	.00	.00	.00
Uncorr'd Factors	448.76	104	<.001	4.31	1343.42	16	<.001	.82	2.73	.12	512.76	.80	.76	.79
One Higher Order Factor	241.01	100	<.001	2.41	207.75	4	<.001	.89	.88	.08	313.01	.92	.90	.92
Corr'd Factors	232.02	98	<.001	2.37	8.99	2	.011	.90	.84	.07	308.02	.92	.90	.92
Males														
Independence	990.91	120	<.001	8.26	--	--	--	.38	4.37	.23	1022.91	.00	.00	.00
Uncorr'd Factors	273.98	104	<.001	2.63	716.94	16	<.001	.80	3.02	.11	337.98	.81	.78	.81
One Higher Order Factor	162.28	100	<.001	1.62	111.69	4	<.001	.87	1.03	.07	234.28	.93	.91	.93
Corr'd Factors	159.84	98	<.001	1.63	2.44	2	.295	.87	1.01	.07	235.84	.93	.91	.93

Table 6
Comparisons Between Measures of Psychological Type

Model	χ^2	df	p	χ^2/df	χ^2 Δ	df Δ	p	GFI	RMR	RMSEA	AIC	IFI	TLI	CFI
Independence	4095.72	190	<.001	21.56	--	--	--	.329	13.4	.229	4135.72	.00	.00	.00
Uncorr'd Factors	789.38	170	<.001	4.64	3306.34	20	<.001	.83	***	.10	869.38	.84	.82	.84
One Higher Order Factor	486.89	166	<.001	2.93	302.49	4	<.001	.88	2.58	.07	574.89	.92	.91	.92
Corr'd Factors	482.27	164	<.001	2.94	4.62	2	.099	.89	2.62	.07	574.27	.92	.91	.92

Figure Caption

<u>Figure 1.</u> Assessment of the multivariate normality of item parcels for the entire population.

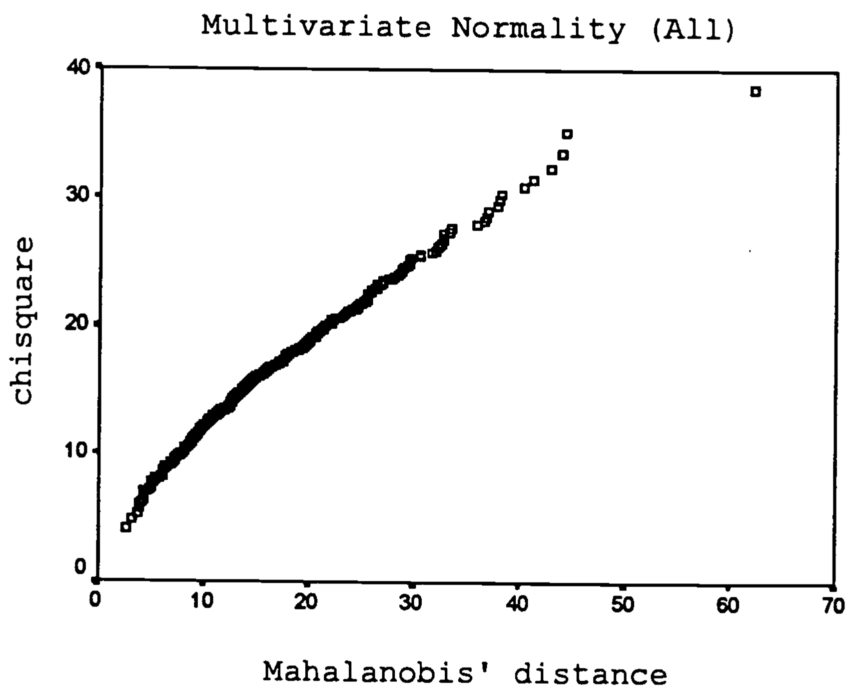
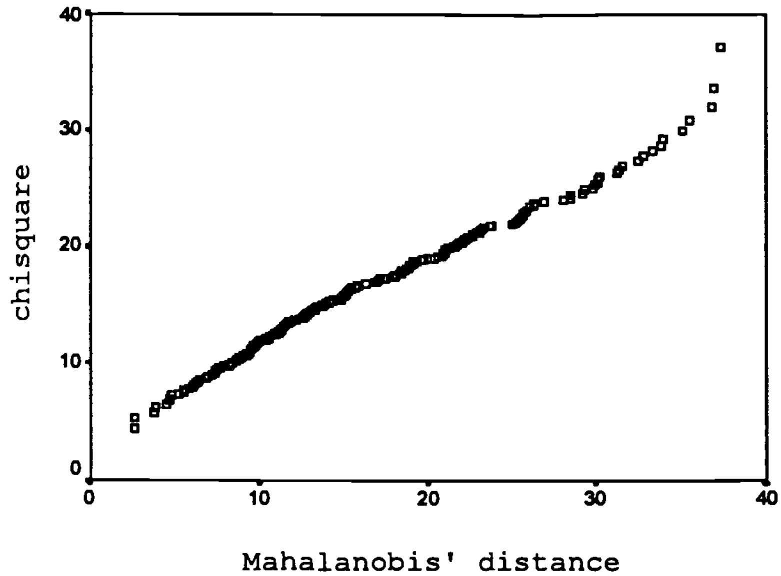

Multivariate Normality (All)
chisquare
Mahalanobis' distance
0
10
20
30
40
0
10
20
30
40
50
60
70

Figure Caption

Figure 2. Assessment of the multivariate normality of item parcels for females and males.

Multivariate Normality (Females)



Multivariate Normality (Males)

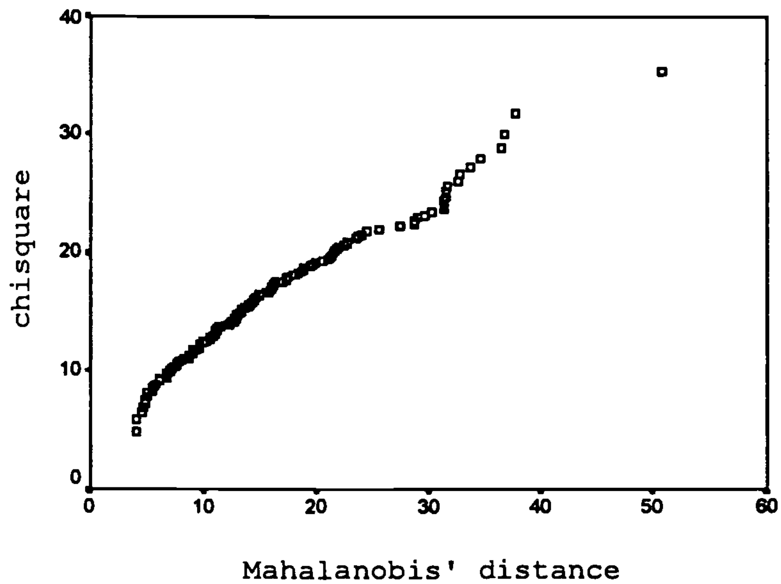
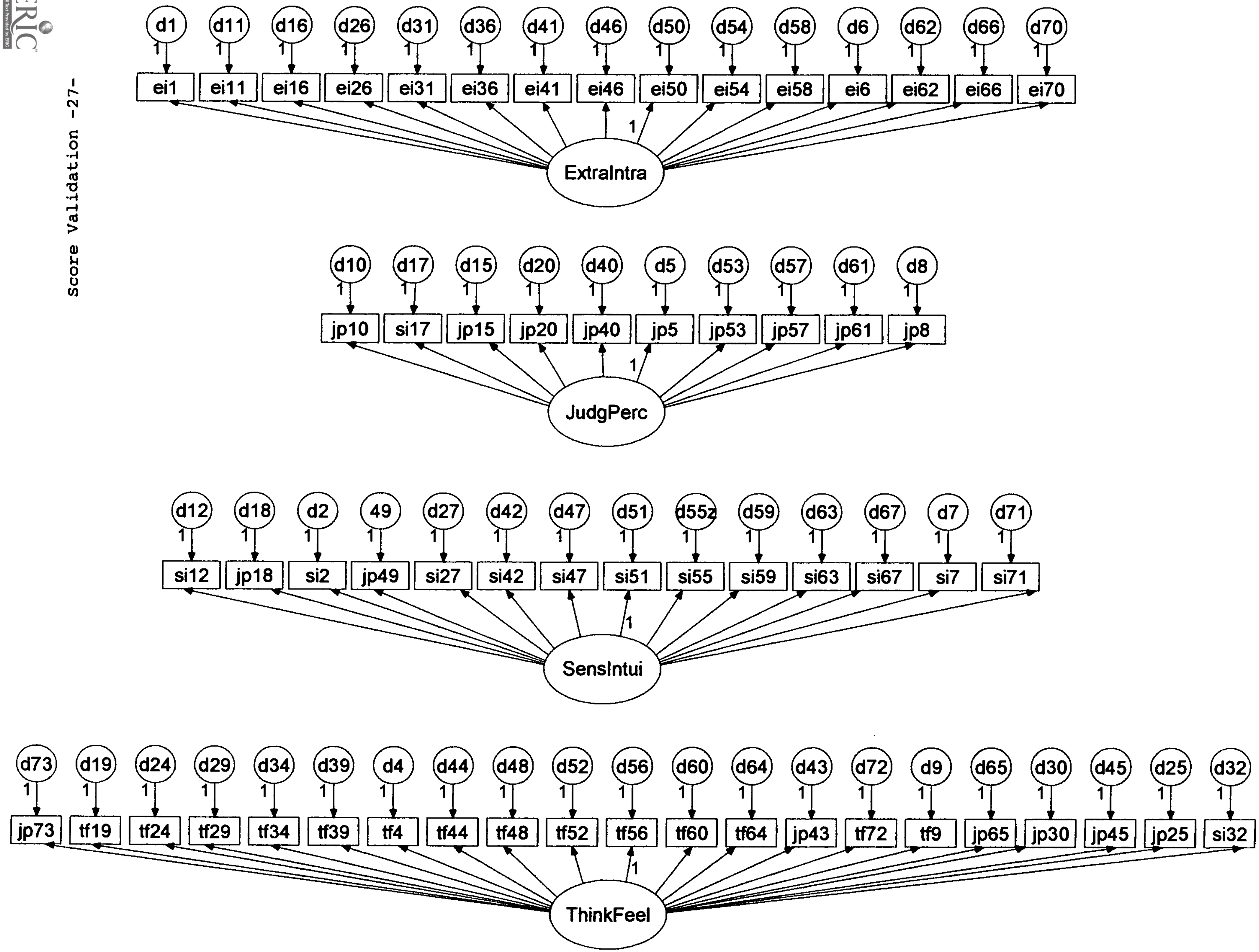


Figure Caption

Figure 3. Model depicting four uncorrelated factors (all items used).


d1 d11 d16 d26 d31 d36 d41 d46 d50 d54 d58 d6 d62 d66 d70
ei1 ei11 ei16 ei26 ei31 ei36 ei41 ei46 ei50 ei54 ei58 ei6 ei62 ei66 ei70
ExtraIntra
d10 d17 d15 d20 d40 d5 d53 d57 d61 d8
jp10 si17 jp15 jp20 jp40 jp5 jp53 jp57 jp61 jp8
JudgPerc
d12 d18 d2 49 d27 d42 d47 d51 d55z d59 d63 d67 d7 d71
si12 jp18 si2 jp49 si27 si42 si47 si51 si55 si59 si63 si67 si7 si71
SensIntui
d73 d19 d24 d29 d34 d39 d4 d44 d48 d52 d56 d60 d64 d43 d72 d9 d65 d30 d45 d25 d32
jp73 tf19 tf24 tf29 tf34 tf39 tf4 tf44 tf48 tf52 tf56 tf60 tf64 jp43 tf72 tf9 jp65 jp30 jp45 jp25 si32
ThinkFeel
1

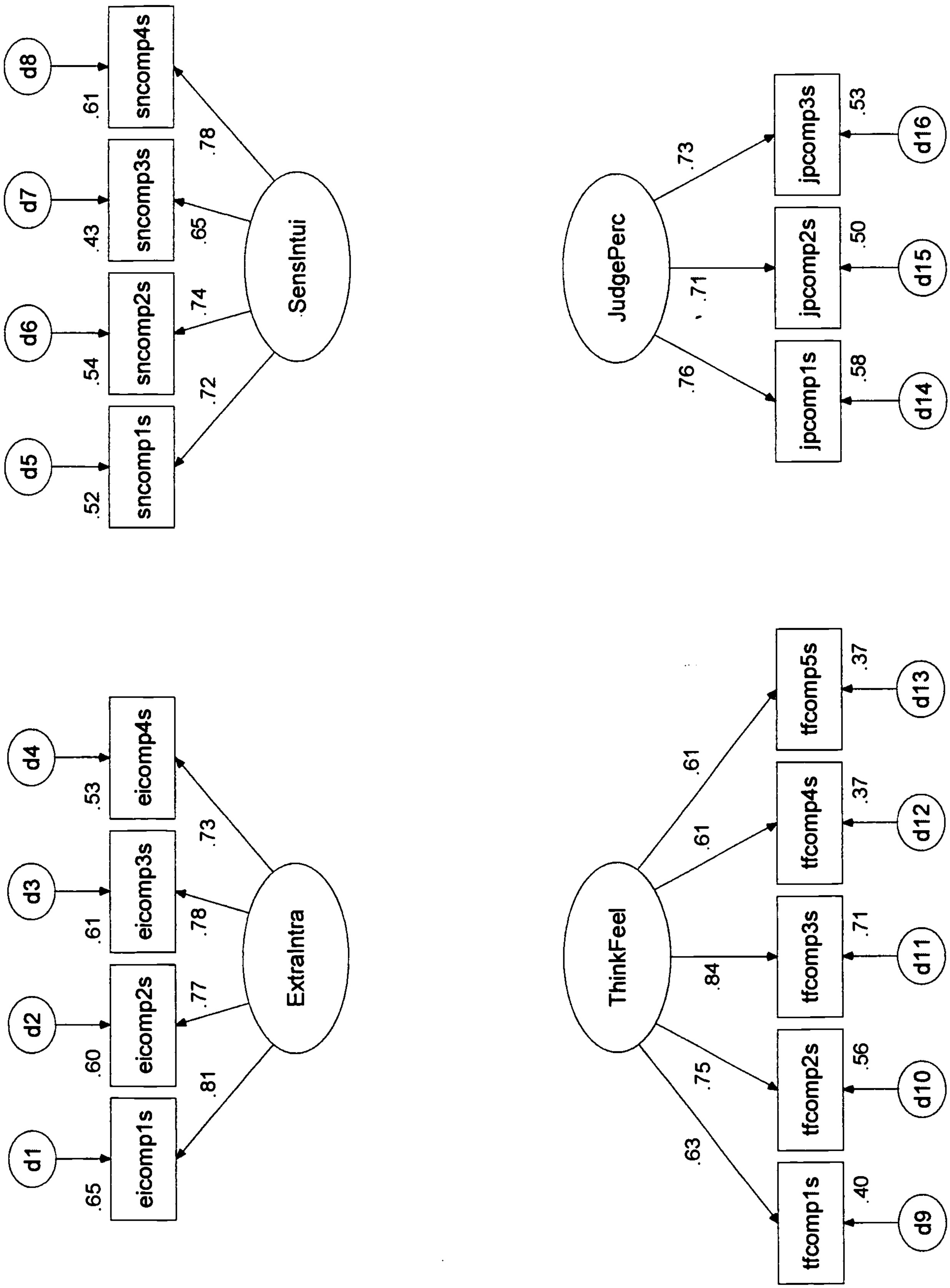
31

32

Figure Caption

Figure 4. Standardized solution for model with four uncorrelated factors (item parcels used).

Note. The values in the graph beside the observed variables are not the error variances, but here are the percentages of variance explained by the model for each observed variable.


ExtraIntra
eicomp1s
eicomp2s
eicomp3s
eicomp4s
d1
d2
d3
d4
.81
.77
.78
.73
.65
.60
.61
.53
SensIntui
sncomp1s
sncomp2s
sncomp3s
sncomp4s
d5
d6
d7
d8
.72
.74
.65
.78
.52
.54
.43
.61
ThinkFeel
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
d9
d10
d11
d12
d13
.63
.75
.84
.61
.61
.40
.56
.71
.37
.37
JudgePerc
jpcomp1s
jpcomp2s
jpcomp3s
d14
d15
d16
.76
.71
.73
.58
.50
.53

34

35

Figure Caption

Figure 5. Standardized solution for model with four correlated factors.

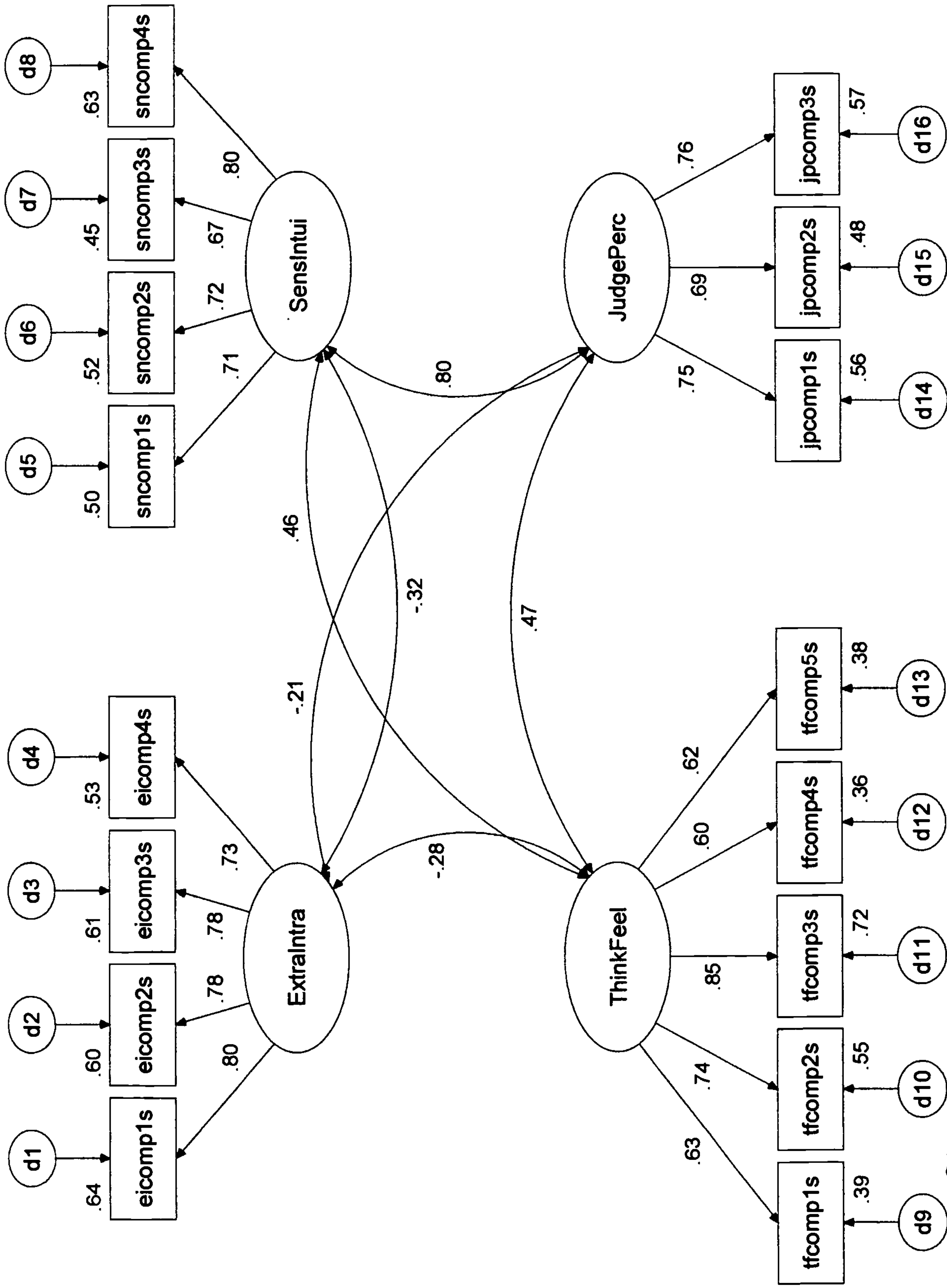

ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
.64
.60
.61
.53
.80
.78
.78
.73
.50
.52
.45
.63
.71
.72
.67
.80
.80
.46
-.32
-.21
.47
-.28
.63
.74
.85
.60
.62
.39
.55
.72
.36
.38
.75
.69
.76
.56
.48
.57

Figure Caption

Figure 6. Standardized solution for model with one general factor.

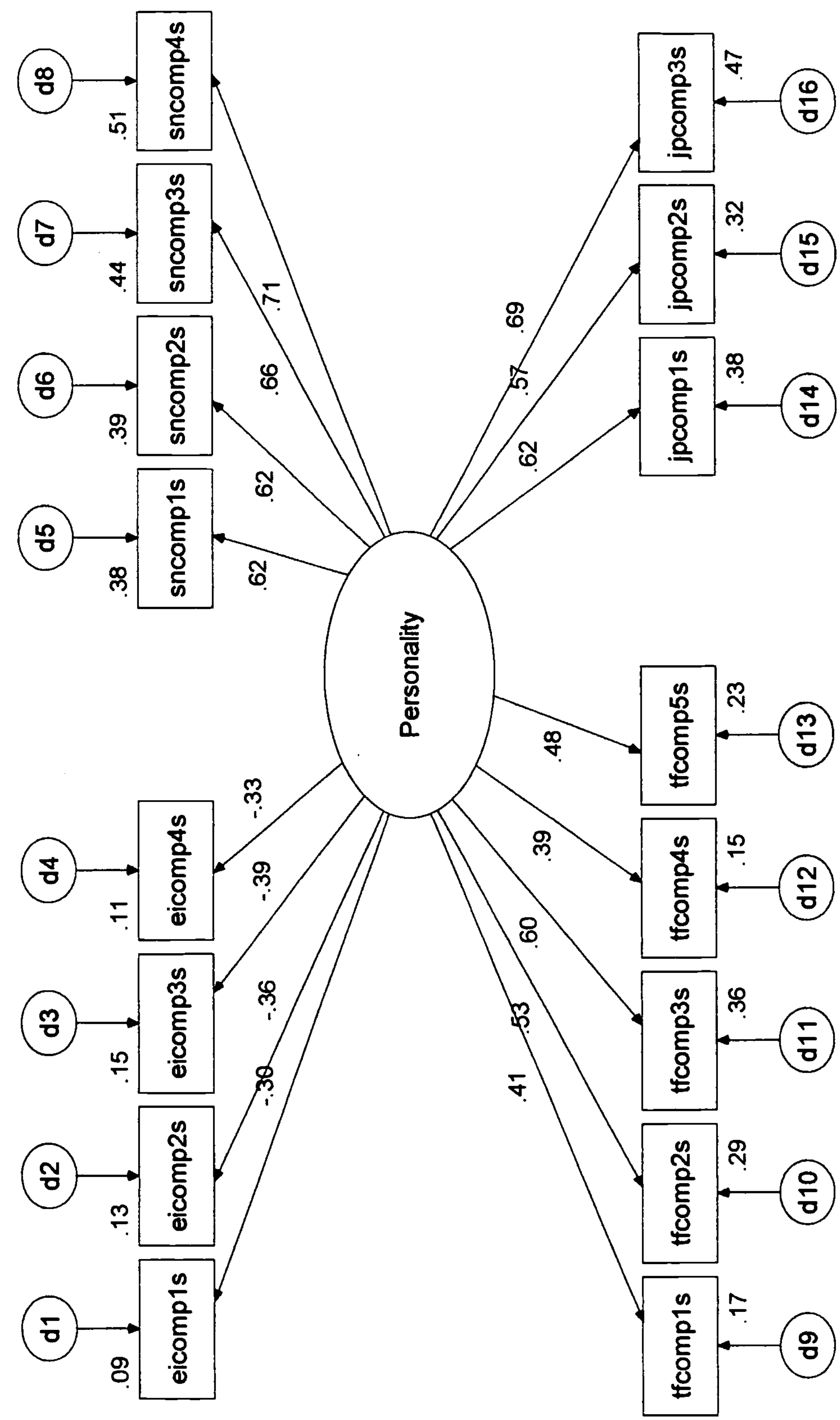

Personality
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
.09
.13
.15
.11
.38
.39
.44
.51
.17
.29
.36
.15
.23
.38
.32
.47
-.30
-.36
-.39
-.33
.62
.62
.66
.71
.41
.53
.60
.39
.48
.62
.57
.69

Figure Caption

Figure 7. Standardized solution for model with three correlated factors.

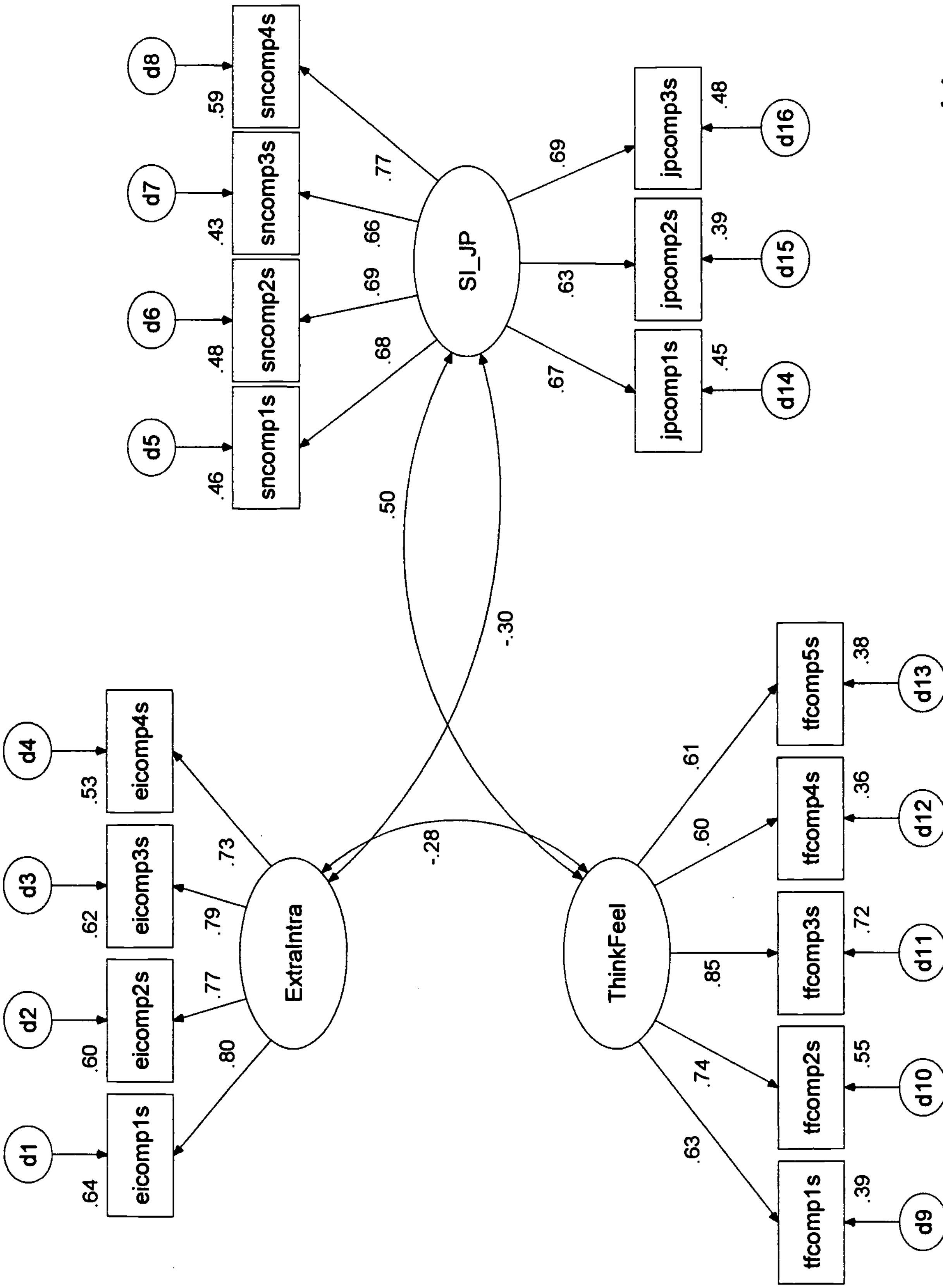

SI_JP
ExtraIntra
ThinkFeel
sncomp1s
sncomp2s
sncomp3s
sncomp4s
jpcomp1s
jpcomp2s
jpcomp3s
eicomp1s
eicomp2s
eicomp3s
eicomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
.46
.48
.43
.59
.68
.69
.66
.77
.67
.63
.69
.45
.39
.48
.64
.60
.62
.53
.80
.77
.79
.73
.63
.74
.85
.60
.61
.39
.55
.72
.36
.38
.50
-.30
-.28

Figure Caption

Figure 8. Standardized solution for model with two higher-order correlated factors.

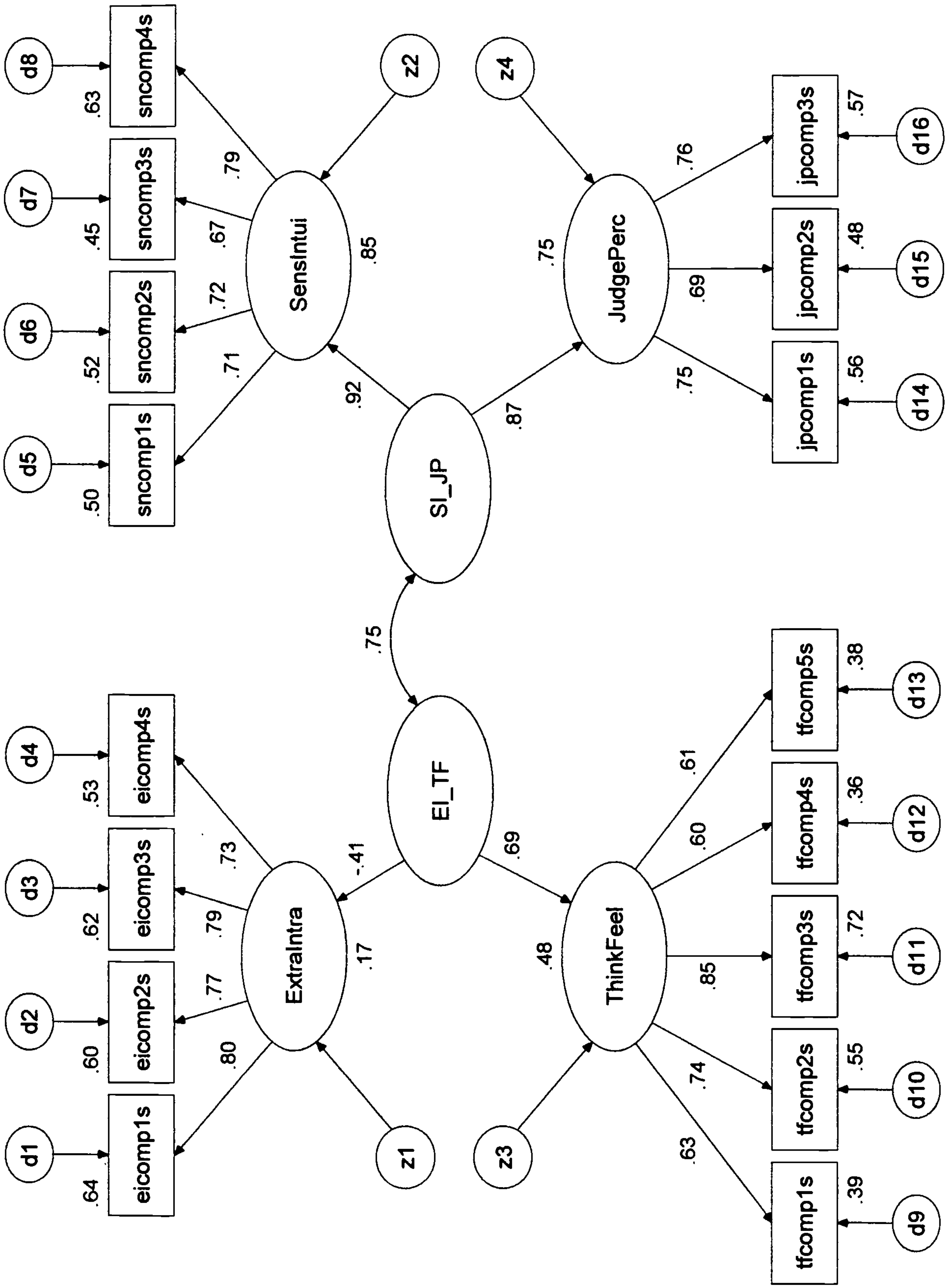

d1
d2
d3
d4
d5
d6
d7
d8
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
.64
.60
.62
.53
.50
.52
.45
.63
.80
.77
.79
.73
.71
.72
.67
.79
ExtraIntra
SensIntui
.17
.85
z1
z2
-.41
.92
EI_TF
SI_JP
.75
.69
.87
z3
z4
.48
.75
ThinkFeel
JudgePerc
.63
.74
.85
.60
.61
.75
.69
.76
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
.39
.55
.72
.36
.38
.56
.48
.57
d9
d10
d11
d12
d13
d14
d15
d16

Figure Caption

Figure 9. Standardized solution for model with one higher-order factor.

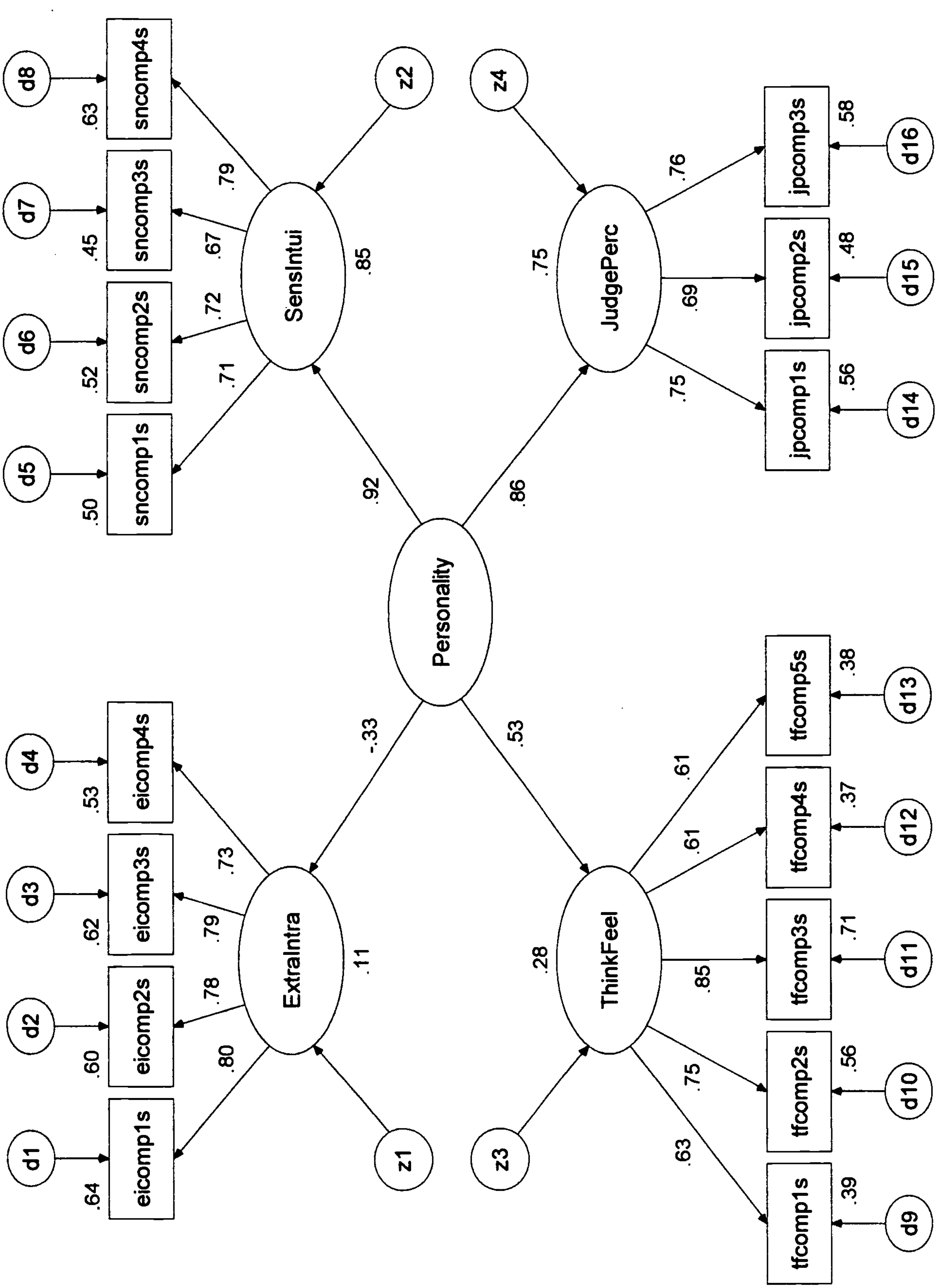

Personality
ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
z1
z2
z3
z4
-.33
.92
.53
.86
.11
.85
.28
.75
.80
.78
.79
.73
.71
.72
.67
.79
.63
.75
.85
.61
.61
.75
.69
.76
.64
.60
.62
.53
.50
.52
.45
.63
.39
.56
.71
.37
.38
.56
.48
.58

Figure Caption

<u>Figure 10.</u> Standardized solution for model with one higher-order factor, first-order to second-order factor paths constrained to be unity.

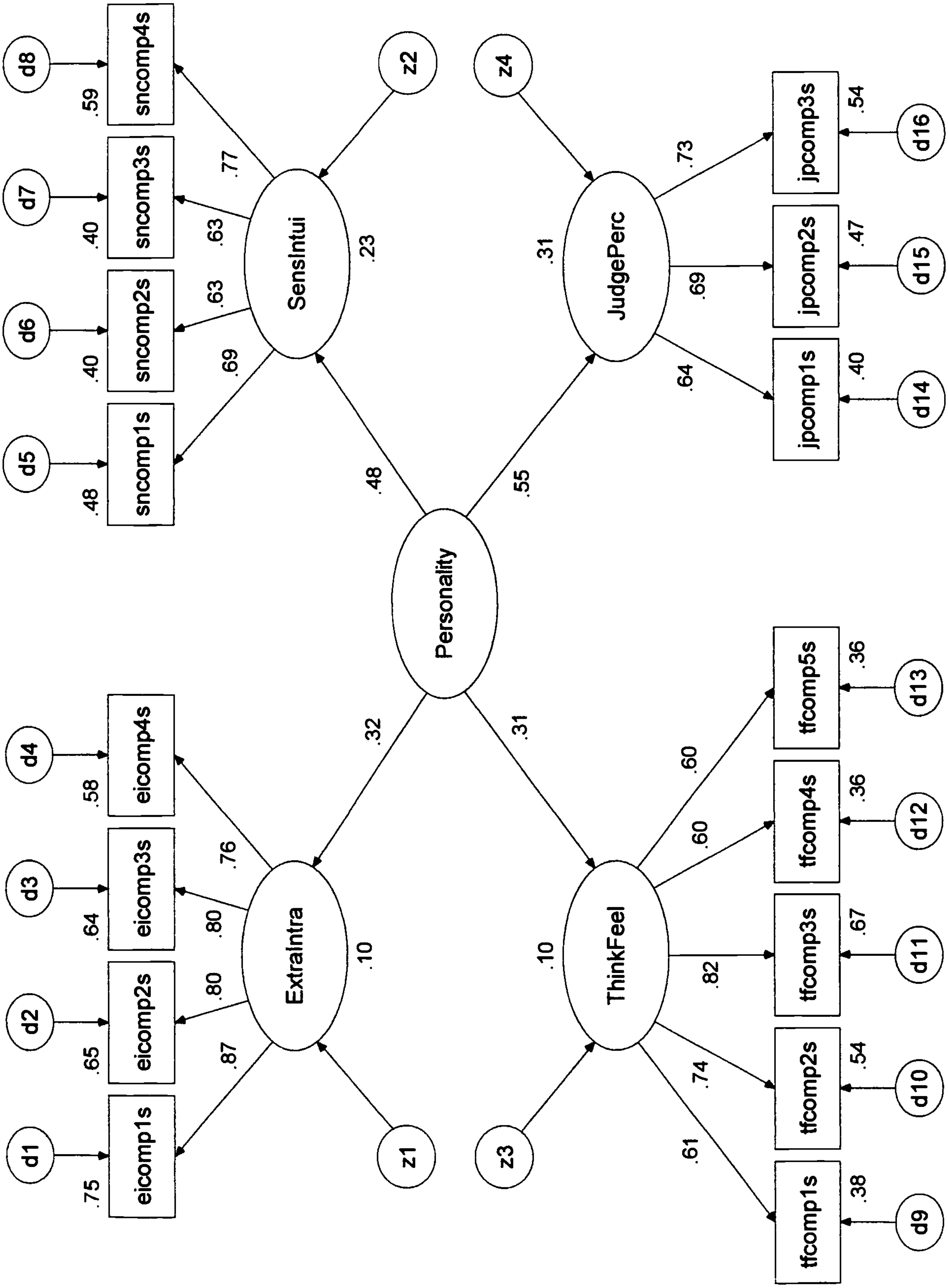

Personality
ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
z1
z2
z3
z4
.32
.48
.31
.55
.10
.23
.10
.31
.87
.80
.80
.76
.69
.63
.63
.77
.61
.74
.82
.60
.60
.64
.69
.73
.75
.65
.64
.58
.48
.40
.40
.59
.38
.54
.67
.36
.36
.40
.47
.54

Figure Caption

<u>Figure 11.</u> Standardized solution for model with four correlated factors, female data only.

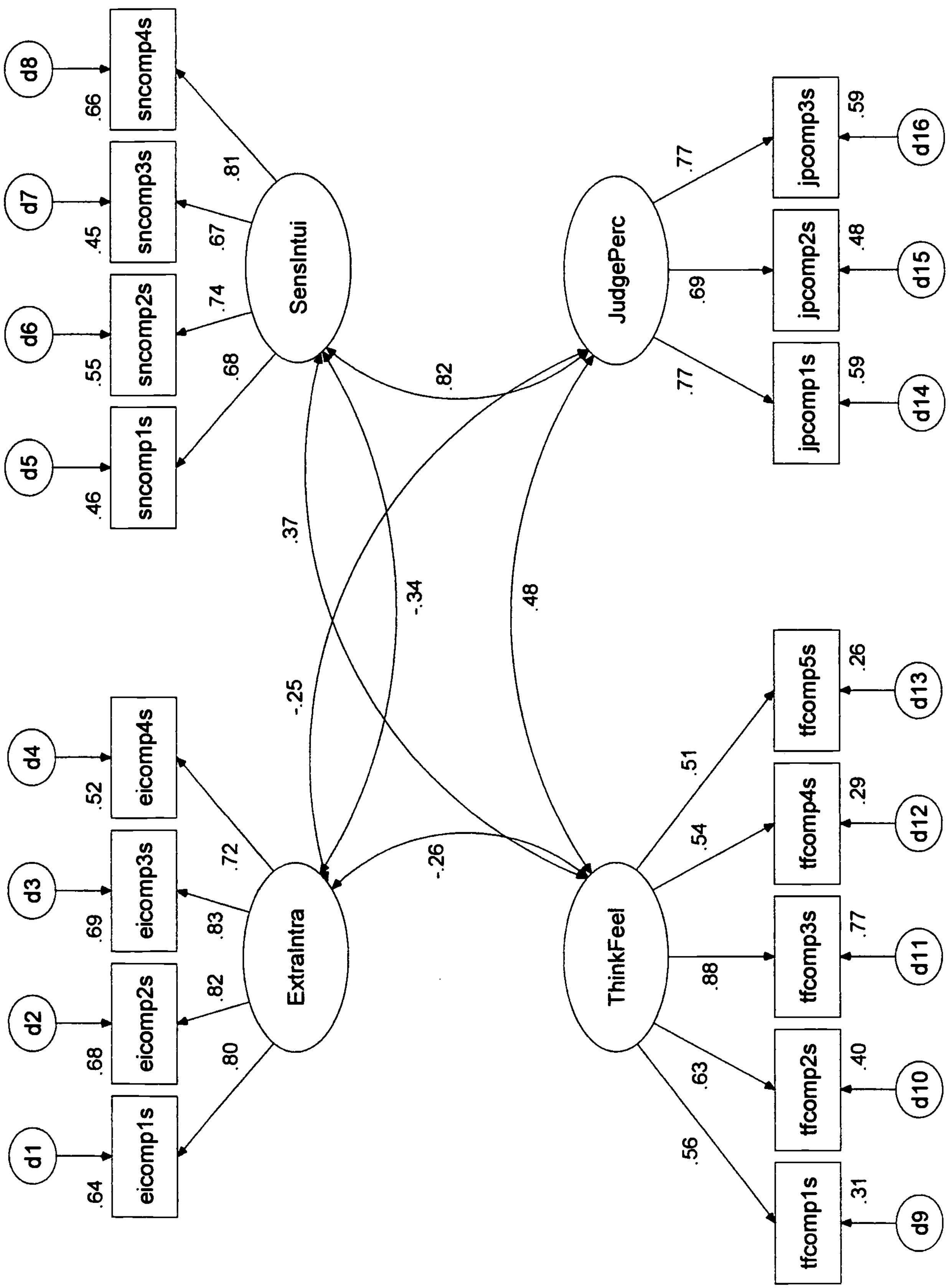

ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
.64
.68
.69
.52
.80
.82
.83
.72
.46
.55
.45
.66
.68
.74
.67
.81
.82
.37
-.34
-.25
.48
-.26
.56
.63
.88
.54
.51
.31
.40
.77
.29
.26
.77
.69
.77
.59
.48
.59

Figure Caption

Figure 12. Standardized solution for model with four correlated factors, male data only.

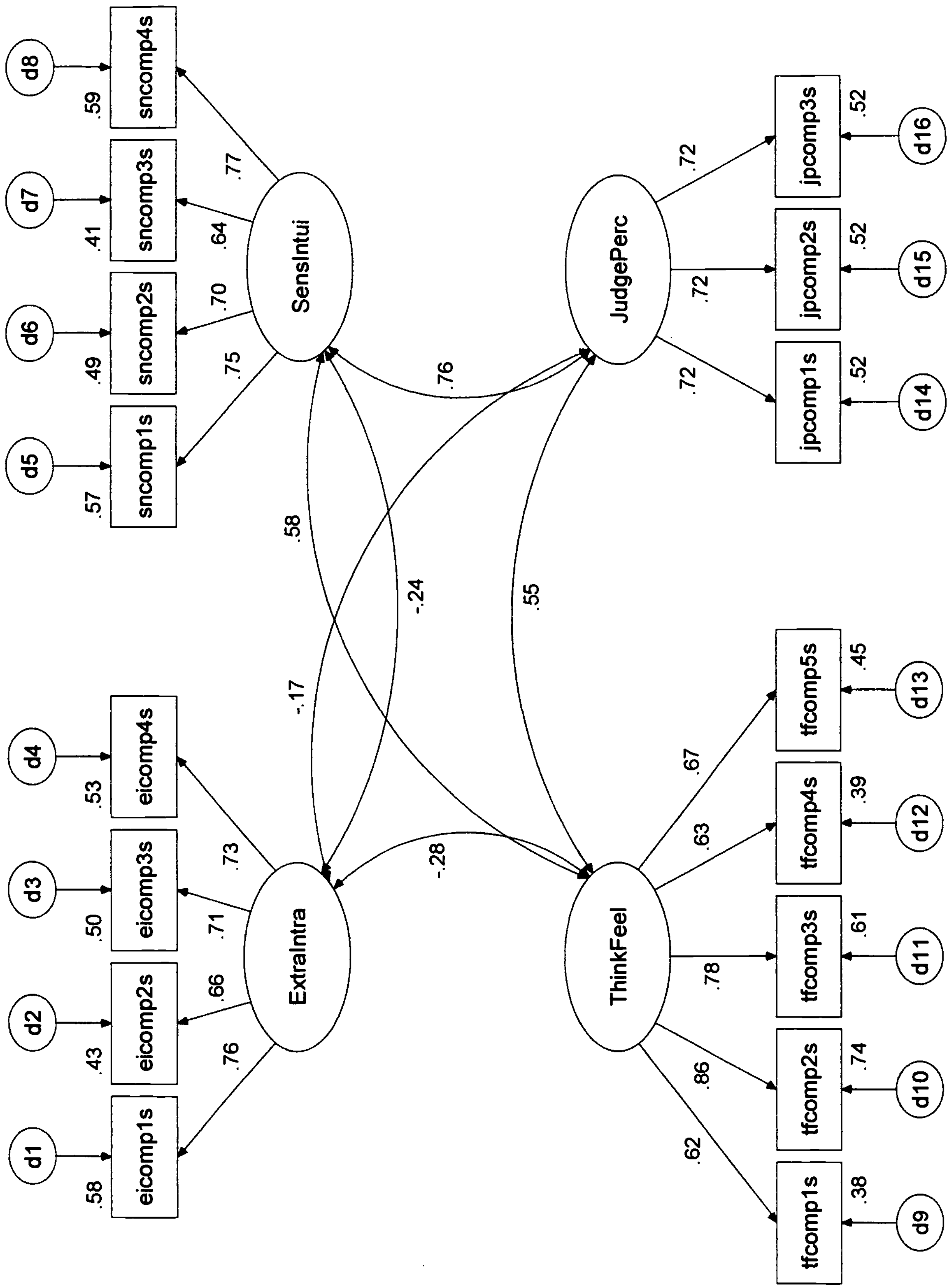

d1
d2
d3
d4
d5
d6
d7
d8
.58
.43
.50
.53
.57
.49
.41
.59
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
.76
.66
.71
.73
.75
.70
.64
.77
ExtraIntra
SensIntui
.58
-.17
-.24
.76
-.28
.55
ThinkFeel
JudgePerc
.62
.86
.78
.63
.67
.72
.72
.72
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
.38
.74
.61
.39
.45
.52
.52
.52
d9
d10
d11
d12
d13
d14
d15
d16

Figure Caption

<u>Figure 13.</u> Standardized solution for model with one higher-order factor, female data only.

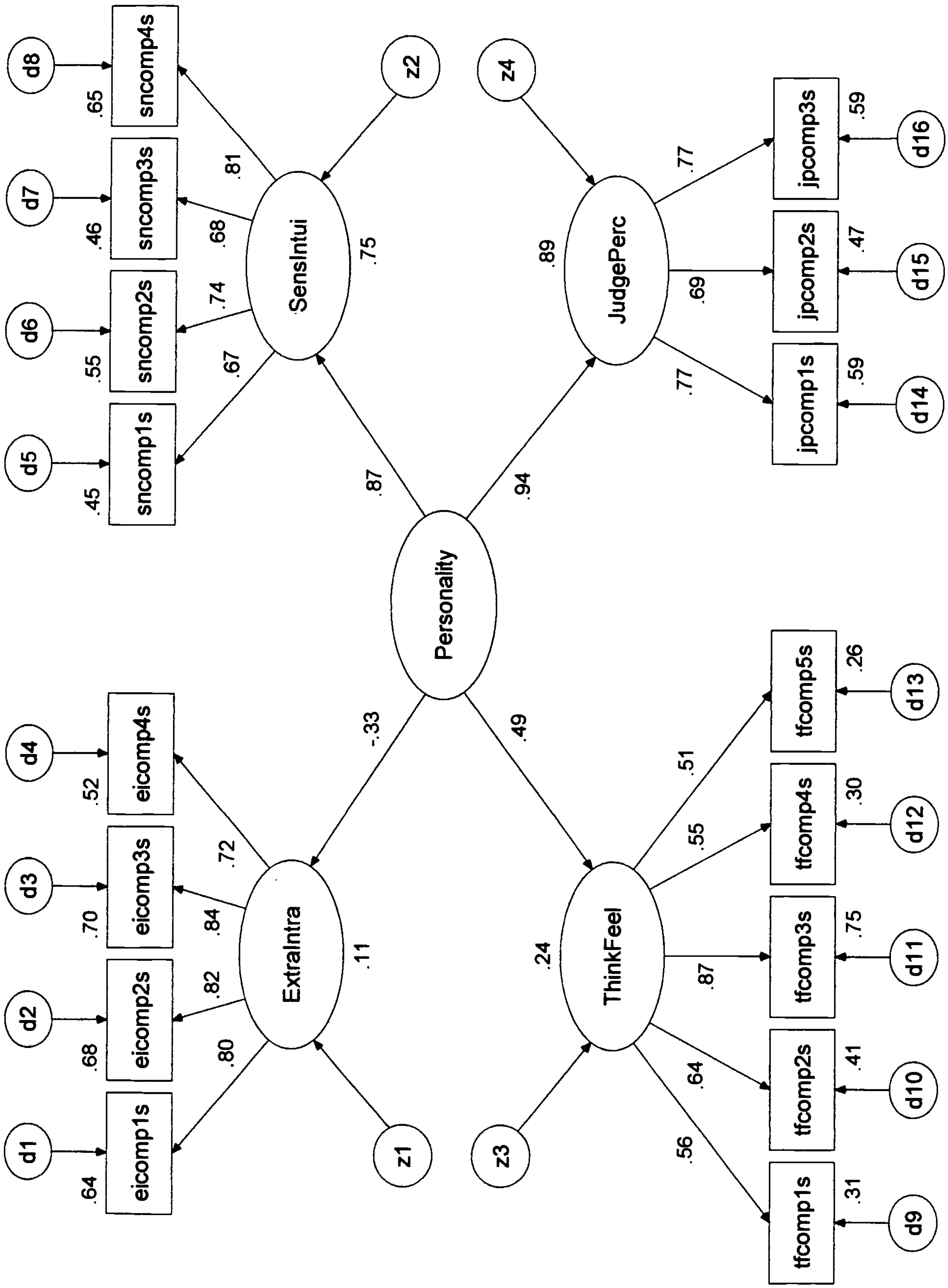
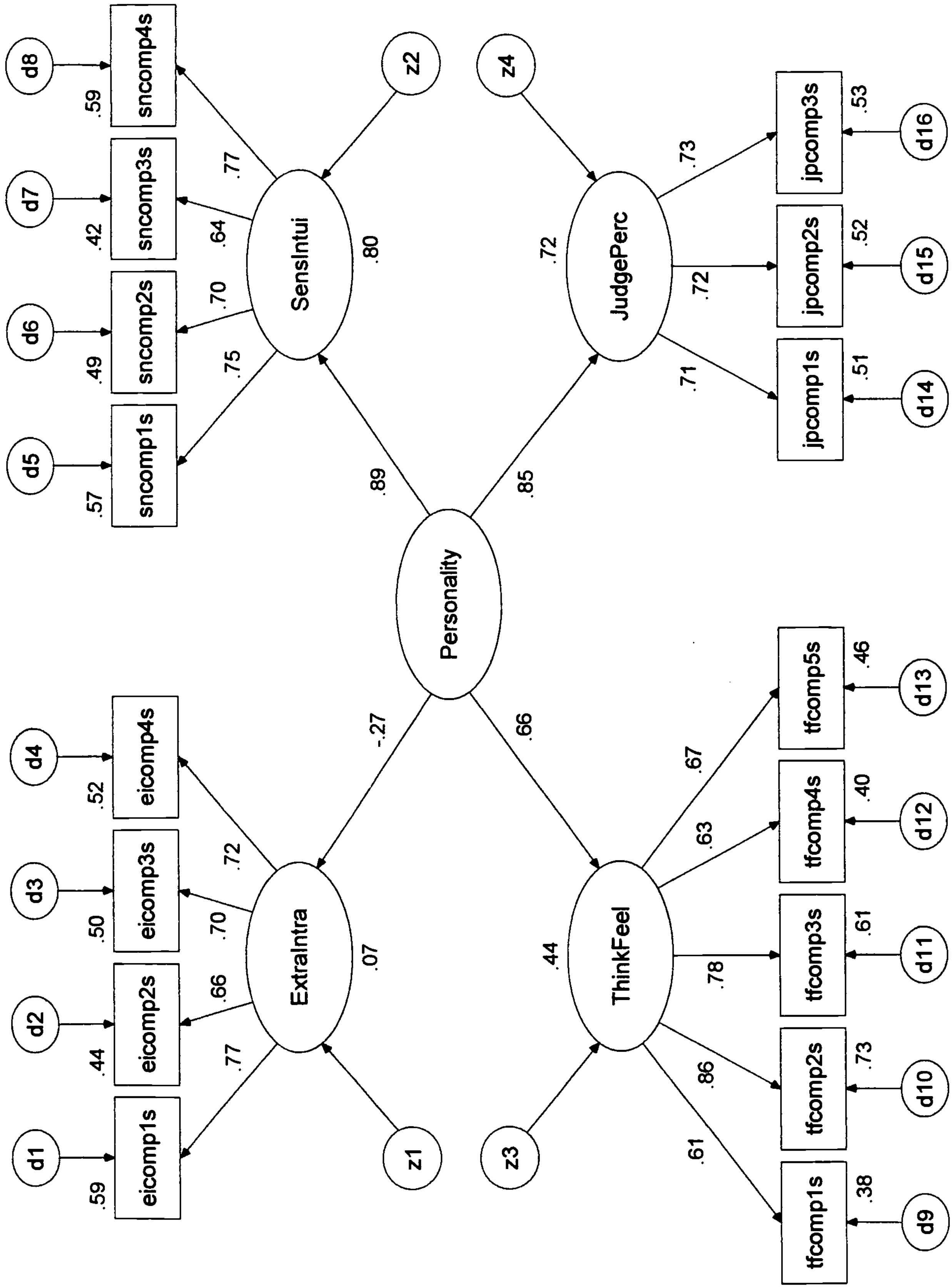

Personality
ExtraIntra
SensIntui
ThinkFeel
JudgePerc
-.33
.87
.49
.94
.11
.75
.24
.89
z1
z2
z3
z4
eicomp1s
eicomp2s
eicomp3s
eicomp4s
.80
.82
.84
.72
.64
.68
.70
.52
d1
d2
d3
d4
sncomp1s
sncomp2s
sncomp3s
sncomp4s
.67
.74
.68
.81
.45
.55
.46
.65
d5
d6
d7
d8
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
.56
.64
.87
.55
.51
.31
.41
.75
.30
.26
d9
d10
d11
d12
d13
jpcomp1s
jpcomp2s
jpcomp3s
.77
.69
.77
.59
.47
.59
d14
d15
d16

Figure Caption

Figure 14. Standardized solution for model with one higher-order factor, male data only.


Personality
ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
sncomp1s
sncomp2s
sncomp3s
sncomp4s
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
jpcomp1s
jpcomp2s
jpcomp3s
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
z1
z2
z3
z4
-.27
.89
.66
.85
.07
.80
.44
.72
.77
.66
.70
.72
.75
.70
.64
.77
.61
.86
.78
.63
.67
.71
.72
.73
.59
.44
.50
.52
.57
.49
.42
.59
.38
.73
.61
.40
.46
.51
.52
.53

64

65

Figure Caption

Figure 15. Standardized solution for model with four correlated factors, PPDSQ and MBTI data combined.

Note. Each MBTI composite is the far right indicator under each factor.

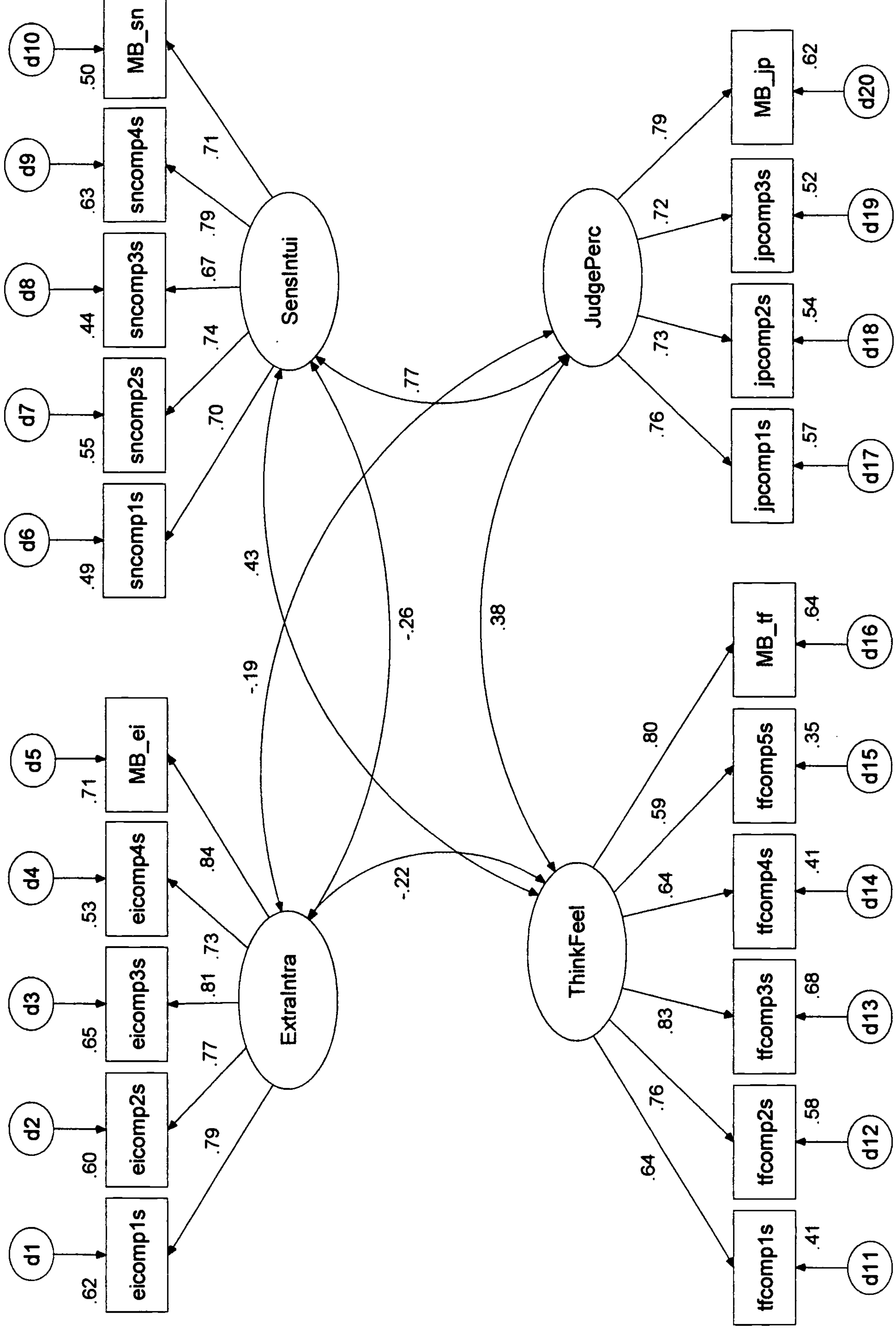
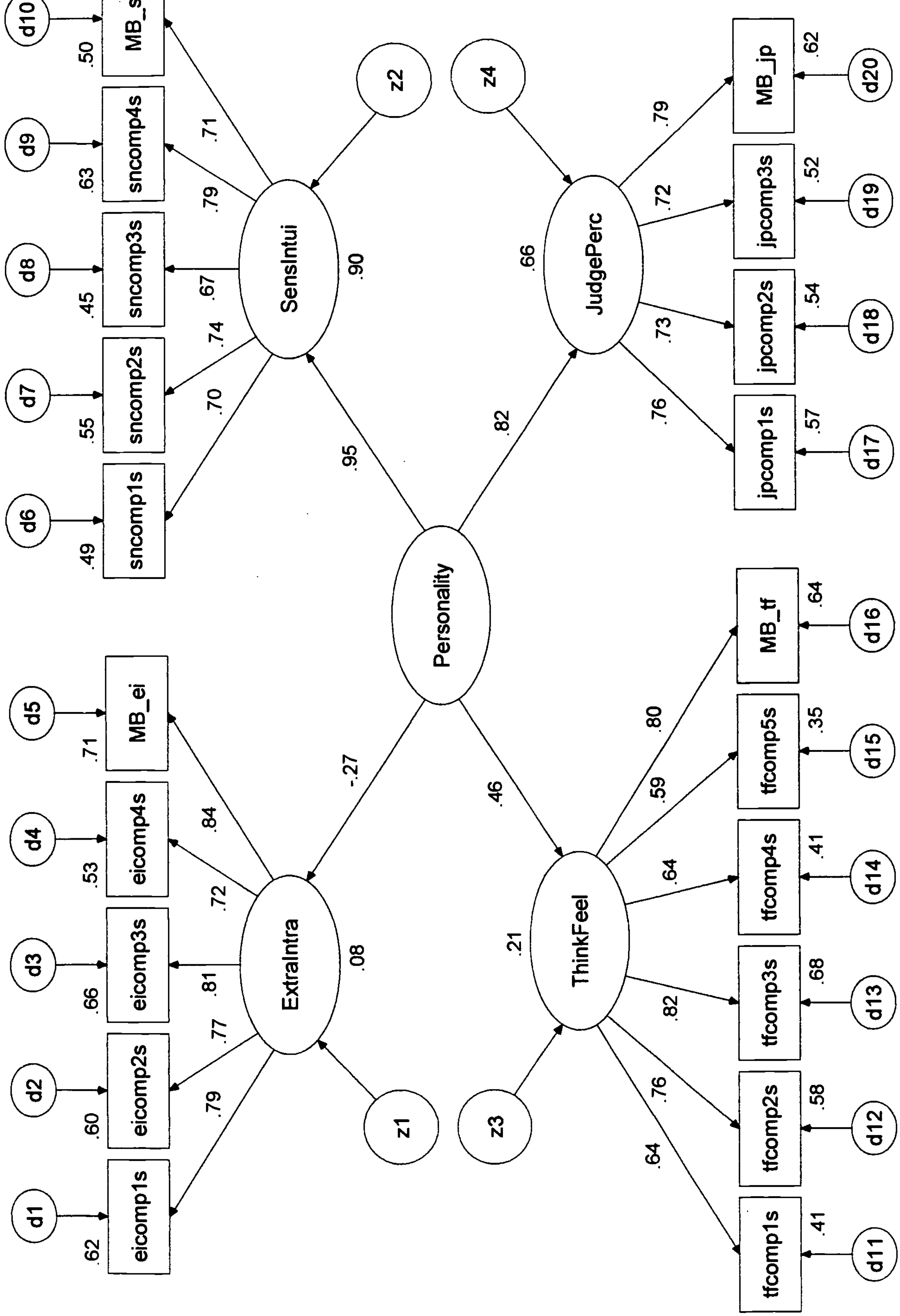

ExtraIntra
SensIntui
ThinkFeel
JudgePerc
eicomp1s
eicomp2s
eicomp3s
eicomp4s
MB_ei
sncomp1s
sncomp2s
sncomp3s
sncomp4s
MB_sn
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
MB_tf
jpcomp1s
jpcomp2s
jpcomp3s
MB_jp
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
d17
d18
d19
d20
.62
.60
.65
.53
.71
.79
.77
.81
.73
.84
.49
.55
.44
.63
.50
.70
.74
.67
.79
.71
.43
-.19
-.26
-.22
.38
.77
.64
.76
.83
.64
.59
.80
.41
.58
.68
.41
.35
.64
.76
.73
.72
.79
.57
.54
.52
.62

Figure Caption

Figure 16. Standardized solution for model with one higher-order factor, PPDSQ and MBTI data combined.


Personality
ExtraIntra
SensIntui
ThinkFeel
JudgePerc
z1
z2
z3
z4
-.27
.95
.46
.82
.08
.90
.21
.66
eicomp1s
eicomp2s
eicomp3s
eicomp4s
MB_ei
.79
.77
.81
.72
.84
.62
.60
.66
.53
.71
d1
d2
d3
d4
d5
sncomp1s
sncomp2s
sncomp3s
sncomp4s
MB_sn
.70
.74
.67
.79
.71
.49
.55
.45
.63
.50
d6
d7
d8
d9
d10
tfcomp1s
tfcomp2s
tfcomp3s
tfcomp4s
tfcomp5s
MB_tf
.64
.76
.82
.64
.59
.80
.41
.58
.68
.41
.35
.64
d11
d12
d13
d14
d15
d16
jpcomp1s
jpcomp2s
jpcomp3s
MB_jp
.76
.73
.72
.79
.57
.54
.52
.62
d17
d18
d19
d20